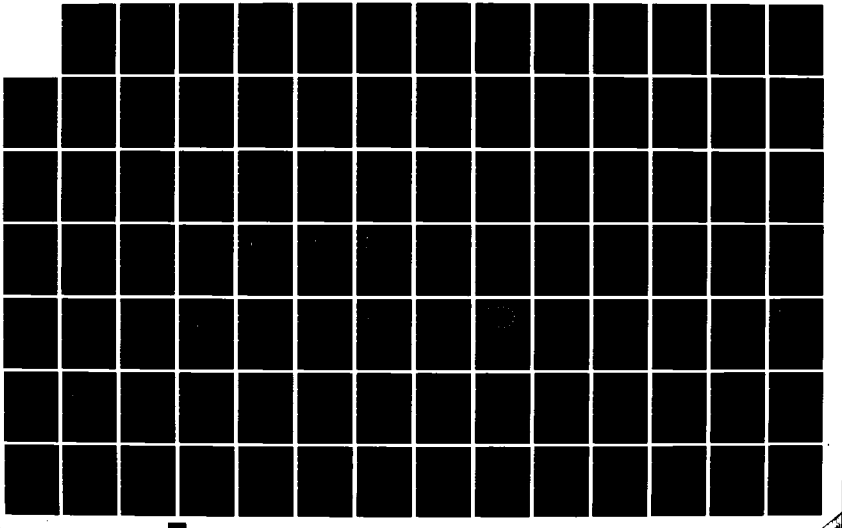
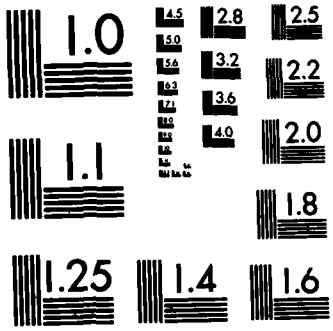
AD-A123 177 ESTIMATING THE PARAMETERS OF EXPONENTIALLY DAMPED OR UNDAMPED SINUSOIDAL.. (U) RHODE ISLAND UNIV KINGSTON DEPT OF ELECTRICAL ENGINEERING R KUMARESAN OCT 82 1/2

UNCLASSIFIED N00014-81-K-0144 F/G 17/2 NL





(12)

Report Number 2

ESTIMATING THE PARAMETERS OF EXPONENTIALLY DAMPED OR UNDAMPED SINUSOIDAL SIGNALS IN NOISE

Ramdas Kumaresan

Department of Electrical Engineering
University of Rhode Island
Kingston, Rhode Island 02881

October 1982

Prepared for

OFFICE OF NAVAL RESEARCH (Code 411SP)
Statistics and Probability Branch
Arlington, Virginia 22217
under Contract N00014-81-K-0144
D. W. Tufts, Principal Investigator

Approved for public release; distribution unlimited

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1 REPORT NUMBER 2	2 GOVT ACCESSION NO. AD-A123 177	3 RECIPIENT'S CATALOG NUMBER
4 TITLE (and Subtitle) ESTIMATING THE PARAMETERS OF EXPONENTIALLY DAMPED OR UNDAMPED SINUSOIDAL SIGNALS IN NOISE		5 TYPE OF REPORT & PERIOD COVERED Technical Report Feb. 1981 - July 1982 6 PERFORMING ORG REPORT NUMBER
7. AUTHOR(s) Ramdas Kumaresan		8. CONTRACT OR GRANT NUMBER(s) N00014-81-K-0144
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Electrical Engineering University of Rhode Island Kingston, Rhode Island 02881		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR SRO-103
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research (Code 411 SP) Department of the Navy Arlington, Virginia 22217		12. REPORT DATE October 1982 13. NUMBER OF PAGES 122
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Also submitted as the Ph.D. Thesis of Ramdas Kumaresan to the Electrical Engineering Department, University of Rhode Island, Kingston, RI 02881

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Prony's Method, Spectral Estimation, High-Resolution Estimation, Linear Prediction, Estimation of Signal Parameters

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

We address the problem of estimating the parameters of a signal embedded in noise. The signal is composed of samples of a sum of M exponentially damped or undamped sinusoidal signals. That is, s(n) the signal samples, are given by the formula $s(n) = \sum_{k=1}^{M} a_k e^{s_k n}$, where $\{a_k\}$ are the amplitudes of the exponentials. The problem is to estimate

DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(Abstract con't.)

the value of M, $\{a_k\}$ and $\{s_k\}$ from a short record of the data samples which are corrupted by noise.

We suggest improvements to the original approach proposed by Prony. He observed that the samples s(n) obey an M th order difference equation and that from the roots of the characteristic polynomial of the difference equation, the parameters $\{s_k\}$ can be determined. We present three methods to improve this approach when the signal is noise corrupted. The two key ideas used in the improvements are listed below in the order of importance.

(1) By using a higher order (greater than M) difference equation formed with the data samples, we try to explain some of the noise in the data by a few additional exponentials, which are subsequently eliminated.

(2) The signal samples s(n) when embedded in a (Toeplitz or Hankel) matrix result in a rank deficient matrix. This is a special property of the exponential signals and is independent of the signal parameters. We make use of this knowledge regarding the signal to improve the signal to noise ratio (SNR) in the data. This is accomplished in a natural way using singular value decomposition (SVD) of a matrix.

The improvement in performance of our methods over traditional methods is demonstrated by computer simulation.

ABSTRACT

We address the problem of estimating the parameters of a signal embedded in noise. The signal is composed of samples of a sum of M exponentially damped or undamped sinusoidal signals. That is, s(n) the signal samples, are given by the formula $s(n) = \sum_{k=1}^{M} a_k e^{s_k n}$, where a_k are the amplitudes of the exponentials. The problem is to estimate the value of M, $\{a_k\}$ and $\{s_k\}$ from a short record of the data samples which are corrupted by noise.

We suggest improvements to the original approach proposed by Prony. He observed that the samples s(n) obey an M th order difference equation and that from the roots of the characteristic polynomial of the difference equation, the parameters $\{s_k\}$ can be determined. We present three methods to improve this approach when the signal is noise corrupted. The two key ideas used in the improvements are listed below in the order of importance.

(1) By using a higher order (greater than M) difference equation formed with the data samples, we try to explain some of the noise in the data by a few additional exponentials, which are subsequently eliminated.

(2) The signal samples s(n) when embedded in a (Toeplitz or Hankel) matrix result in a rank defficient matrix. This is a special property of the exponential signals and is independent of the signal parameters. We make use of this knowledge regarding the signal to improve the signal to noise ratio (SNR) in the data. This is accomplished in a natural way using singular value decomposition (SVD) of a matrix.

The improvement in performance of our methods over traditional methods is demonstrated with computer simulation.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONT.)

CHAPTER 1: INTRODUCTION

1.1: PROBLEM STATEMENT:

We address the problem of estimating certain unknown parameters of signals corrupted by noise. Let us assume that a sequence of numbers, $y(1),y(2),\ldots y(N)$, which are samples of a noise corrupted signal, is available for processing by a digital computer. The sequence $y(n)$, $n=1,2\ldots N$ is further assumed to be composed of samples of a sum of M unknown, exponentially damped or undamped sinusoidal signals and additive noise $w(n)$. That is,

$$y(n) = \sum_{k=1}^{M} a_k e^{s_k n} + w(n) \quad , n=1,2\ldots N \qquad 1.1$$

The sampling period is assumed to be one second without loss of generality. The Nyquist interval is $0-2\pi$, a_k and s_k, $k=1,2,\ldots M$ are unknown complex numbers, in general. M is also unknown. We write the complex numbers s_k explicitly as follows:

$$s_k = \alpha_k + jw_k$$

The α_k's shall be called the damping factors, and the w_k's are the radian frequencies that will lie between 0 and 2π. j is the square root of minus one.

Given the N samples of noise corrupted data, the problem is to estimate the unknown values of M, a_k, s_k, $k=1,2,\ldots M$.

1.2: SIGNIFICANCE OF THE PROBLEM:

The above is an age old problem and has occurred sometimes in disguise in several areas of applied science. Proliferation of high speed computers has rekindled interest in this problem in the past two decades. The following examples drawn from different areas are representative of the occurrence of this problem.

In speech signal processing literature, it has been observed that the human vocal tract coupled with the nasal cavity behaves approximately like a pole-zero filter [1]. Wideband excitation to this filter is provided by the vibration of the vocal cord in the form of quasi-periodic train of short time duration pulses. Therefore, to a first order approximation [2], each pitch frame (the time duration in between two pitch pulses) of the speech waveform for (nasalized) vowel sounds can be assumed to be a sum of exponentially damped sinusoidal signals. The sequence of noise values, w(n), in formula 1.1, represents the error in this assumption plus any observation noise. Thus, given a sequence of samples of the speech waveform from individual pitch frames, one is interested in obtaining a pole-zero representation for the vocal tract, for speech data compression purposes.

In another example the signals could be undamped sinusoidal signals, closely spaced in frequency compared to the reciprocal of the observation interval of N seconds. This corresponds to assuming $\alpha_k = 0$, k=1,2,...M in formula 1.1. This is the so called spectral resolution problem and has received considerable attention in signal processing and statistical time series analysis literature. See for example, the review article by Kay and Marple [3]. Undamped sinusoidal signals occur in a wide variety of practical problems, ranging from resolving the locations of distant stars by a radio telescope [4] to measuring the depth of the ionospheric layer by 'ionospheric sounding' [5]. In these problems a critical factor is the limitation on the number of samples available

for processing. For example, the aperture of the radio telescope, which corresponds to N, is necessarily limited by cost considerations; whereas in the second example, the ionospheric layer may be changing in depth with respect to time, thus limiting the observation interval N.

The third case of interest is when the w_k's in formula 1.1 are zero; that is, the signal is composed of a sum of M exponential decays. Such signals are often observed in biomedical problems involving radioactive tracer data analysis [6]. See also ref. [7].

In all the above examples, it is appropriate to consider the sequence y(n) as a sum of M deterministic signals and white noise based on the physics of signal generation and measurement. However, the model in formula 1.1 can also be used for estimating the spectrum of an autoregressive (AR) or autoregressive moving average (ARMA) process. Such processes are generated when an all pole (AR) or a pole-zero (ARMA) filter is excited by a white noise sequence. Given a short record of the random process, the problem is to estimate the spectrum of the process. This problem can be put in our framework by observing that the true autocorrelation sequence of AR/ARMA processes are samples of a sum of M exponentially damped signals, M being the number of poles in the generating filter [8,9]. Therefore, if we estimate the autocorrelation sequence from the given data, it is reasonable to model the estimated autocorrelation sequence as in formula 1.1. Now, w(n) stands for the difference between the estimated and the true autocorrelation sequence of the process. Hence, if we estimate the unknown parameters, a_k, s_k, and M, it is possible to obtain a spectral estimate of the process by estimating the generating filter transfer function.

1.3: PREVIOUS RESULTS:

Since our problem is an old one (see for example ref. [10]) an exhaustive review of the extensive literature is almost impossible. However, there are at

least three major approaches to estimating the unknown parameters from the N given samples of data y(1),y(2),...y(N). We shall briefly review these below. Some other possible approaches, which are not discussed here, include data extrapolation techniques [11,12] and system identification techniques [13].

1.3.1: DIRECT APPROACH:

In this approach one attempts to fit a signal model to the data sequence y(n) in the least square sense. That is, the error E

$$E = \sum_{n=1}^{N} |y(n) - \hat{h}(n)|^2 \qquad 1.2$$

where $\hat{h}(n) = \sum_{k=1}^{\hat{M}} \hat{a}_k e^{\hat{s}_k n}$, is minimized by choosing $\hat{a}_k$'s and $\hat{s}_k$'s. $\hat{M}$ is assumed number of signal components. If we set the derivatives of E with respect to the unknown parameters to zero in attempting to minimize E, we realize that some of the necessary conditions are in the form of nonlinear equations. Nevertheless, these equations can be solved by using standard nonlinear minimization techniques [14]. See also ref. [15].

A simplification to this approach was suggested by several authors [2,7, 16-18] based on the following observation. If the $\hat{s}_k$'s are known, then the minimization of E is a linear least square problem, by which we mean that the unknown $\hat{a}_k$'s can be found (to minimize E) by solving a set of necessary conditions which are a set of linear equations. Therefore, one can choose values for the $\hat{s}_k$'s, based on some prior knowledge or based on some preprocessing and then find the associated $\hat{a}_k$'s that minimize E. Then the value of $\hat{s}_k$'s can be altered to find new values for $\hat{a}_k$'s that give a lower minimum for E. This process is repeated until a local minimum for E is found. The corresponding values of $\hat{a}_k$'s and $\hat{s}_k$'s are the estimates of the signal parameters. This is essentially a search procedure. If the value of M is not known, then the above procedure can be repeated at different assumed values for $\hat{M}$, starting from $\hat{M}$=1, until the minimum

value of E does not change appreciably.

This approach is optimum in the maximum likelihood (ML) sense if the perturbations, w(n), are independent and Gaussian. But this method has certain disadvantages. Much computation is involved, especially if the true value of M is large and unknown. Also, one could terminate the search in a local minimum instead of the global minimum.

However this approach is not unreasonable in the special case when the data is known to consist of undamped sinewaves and noise [18]. In this case, $\hat{s}_k$'s (=j $\hat{w}_k$'s) are purely imaginary numbers. Therefore, one needs to search for the minimum of E only in the interval 0-2π in $\hat{M}$ dimensions. Also, if the frequencies of the sinewaves are spaced greater than 1/N Hz apart, then one can obtain frequency estimates ($\hat{w}_k$'s) close to optimum (in the ML sense) by computing the discrete Fourier transform (DFT) of the data samples [19]. This can be done efficiently, using the fast Fourier transform algorithm [20]. Other relevant references, where similar direct approaches are used, include McDonough and Huggins [21], Schweppe [22], Van den Boss [23], Tuttle [102], and Aigrain and Williams [103].

1.3.1: ITERATIVE APPROACHES:

Many nonlinear minimization problems can be solved iteratively, by solving a linear minimization problem at every step in the iteration. Our problem is no exception to this rule. We shall briefly review two iterative procedures in this section. Both attempt to minimize the error E in formula 1.2 iteratively.

The first method that we shall briefly describe is that of Steiglitz and McBride [24] and is an out growth of an idea proposed by Kalman [25]. In this method the error E in formula 1.2 is rewritten as follows:

$$E = \sum_{n=0}^{N-1} \left| y(n+1) - \hat{H}(z)\{\delta(n)\} \right|^2, \qquad 1.3$$

where

$$\delta(n) = 1 \qquad n = 0$$
$$= 0 \qquad n \neq 0$$

$\hat{H}(z)\{\delta(n)\}$ reads $\hat{H}(z)$ operating on $\delta(n)$. For example $z^{-1}\{y(n)\} = y(n-1)$, and if $\hat{H}(z) = \hat{h}_0 + \hat{h}_1 z^{-1} + \hat{h}_2 z^{-2}$, then, $\hat{H}(z)\{y(n)\} = \hat{h}_0 + \hat{h}_1\, y(n-1) + \hat{h}_2 y(n-2)$. Formula 1.3 follows from the fact that the signal model $\hat{h}(n) = \sum_{k=1}^{\hat{M}} a_k\, e^{\hat{s}kn}$ can be considered as an impulse response [20] of a discrete time pole-zero filter with transfer function H(z). Then $\hat{h}(n)$ can be written in operator notation [20] as $\hat{h}(n) = \hat{H}(z)\{\delta(n)\} = \frac{\hat{N}(z)}{\hat{D}(z)}\,\delta(n)$, where

$$\hat{H}(z) = \frac{\hat{N}(z)}{\hat{D})z)} = \frac{\sum_{i=0}^{\hat{M}-1} \hat{n}_i\, z^{-i}}{1 + \sum_{i=1}^{\hat{M}} \hat{d}_i\, z^{-i}}$$

Therefore one wishes to minimize E in 1.3 by choosing $\hat{N}(z)/\hat{D}(z)$ or equivalently by finding $\hat{n}_0, \hat{n}_1, \ldots \hat{n}_{M-1}$ and $\hat{d}_1, \hat{d}_2, \ldots \hat{d}_M$. Again, this problem is nonlinear in parameters $\{\hat{d}_k\}$.

Kalman's suggestion was to minimize the following error.

$$\min_{\hat{D}(z),\, \hat{N}(z)} \sum_{n=0}^{N-1} \left|\hat{D}(z)\{y(n+1)\} - \hat{N}(z)\{\delta(n)\}\right|^2 \qquad 1.4$$

This is not the same error as E in 1.3. But the advantage of this modified error criterion is that it can be minimized with respect to the coefficients of the polynomial $\hat{N}(z)$ and $\hat{D}(z)$ by solving a linear least squares problem. Stieglitz and McBride improved on this idea by minimizing the following error iteratively. In the i th iteration, they minimized

$$\min_{\hat{D}_i(z),\, \hat{N}_i(z)} \sum_{n=0}^{N-1} \left|\frac{\hat{D}_i(z)\{y(n+1)\}}{\hat{D}_{i-1}(z)} - \frac{\hat{N}_i(z)\{\delta(n)\}}{\hat{D}_{i-1}(z)}\right|^2 \qquad 1.5$$

with respect to $\hat{D}_i(z)$ and $\hat{N}_i(z)$, with $\hat{D}_{i-1}(z)$ known from the previous iteration.

To start the iteration $\hat{D}_0(z)$ is calculated by some simple procedure. The work involved in each iteration is the same as in Kalman's method; i.e. only linear least squares problems need be solved However the data sequence y(n) is prefiltered at every iteration (i.e. $y(n)/\hat{D}_{i-1}(z)$ is used instead of y(n)). If $\hat{D}_i(z)$ converges to $\hat{D}(z)$, the above error coincides with the desired error in formula 1.3

This method may have convergence problems [26], especially when the data consists of slightly damped or undamped sinusoidal signals in noise. Also, there seems to be no systematic procedure for selecting the order of the numerator and denominator polynomials. See Done and Rushforth [27], Kay [28], and Stoica and Soderstorm [29] for results related to this method.

The next method we briefly allude to is that of Evans and Fischl [30]. In this method the authors make use of the following observation regarding the error in formula 1.3. For a given polynomial $\hat{D}(z)$, E can be minimized by solving a linear least square problem by finding the numerator polynomial coefficients. Therefore the error E which is a function of $\hat{D}(z)$ and $\hat{N}(z)$ can be rewritten as an explicit function of $\hat{D}(z)$ only. That is, $E(\hat{N}(z),\hat{D}(z))=E_1(\hat{D}(z))$. Then an iterative procedure is used to minimize $E_1(\hat{D}(z))$ with respect to $\hat{D}(z)$. This method resembles the search procedure described in section 1.3.1 where the minimization was carried out with respect to the parameters $\{s_k\}$. Although this technique was intended for a filter design problem, it can be also used for our estimation problem. The difficulties associated with this method include order selection for the polynomials and computational cost.

1.3.3: INDIRECT APPROACH:

This approach is based on an idea proposed by Prony [31] in 1795. His idea was to transform the nonlinear estimation problem into a linear one. It

is based on the following observation. For example, let us assume that the data sequence y(n) consists of two unknown exponential signals only, i.e. $y(n)=\sum_{k=1} a_k e^{s_k n}$. We want to determine the values of s_1, s_2, a_1, and a_2. Note that a polynomial $G(z)=1+g_1 z^{-1}+g_2 z^{-2}$, with $g_1=-(e^{s_1}+e^{s_2})$ and $g_2=e^{s_1+s_2}$ 'annihilates' the sequence y(n). By annihilates [20] we mean that G(z){y(n)} will equal zero for all n. This polynomial is termed 'prediction-error polynomial' here, due to its close connection with the linear prediction analysis [39]. Note that the polynomial G(z) has roots at e^{s_1} and e^{s_2}. Therefore if we can find the coefficients of the polynomial G(z), namely g_1 and g_2, given the sequence y(n), we can find s_1 and s_2 by finding the roots (zeros) of the polynomial. The coefficients of the polynomial can be found by solving the following equations.

$$\begin{bmatrix} y(3) & y(2) & y(1) \\ y(4) & y(3) & y(2) \end{bmatrix} \begin{bmatrix} 1 \\ g_1 \\ g_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad 1.6$$

The right hand side of the above equations is set to zero because G(z) annihilates the sequence y(n). Ideally, only 4 (two times the number of exponential signals in data) contiguous samples of y(n) are needed to find g_1 and g_2. Once s_1 and s_2 are found, the values of a_1 and a_2 can be found by solving the following linear equations.

$$\begin{bmatrix} e^{s_1} & e^{s_2} \\ e^{2s_1} & e^{2s_2} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} y(1) \\ y(2) \end{bmatrix} \qquad 1.7$$

Summarizing, Prony's contribution was to transform the problem of finding the nonlinearly entering parameters s_k, k=1,2,... M into a problem of finding the coefficients $g_1, g_2, \ldots . g_M$ of an M th degree polynomial $G(z)=1+\sum_{k=1}^{M} g_M z^{-k}$. The degree of the polynomial G(z) should equal or exceed M, the number of signal components in the data. $\{g_k\}$ are found by solving linear equations. The $\{s_k\}$ are found from the zero locations of G(z). The amplitudes $\{a_k\}$ are subsequently found by solving yet another system of linear equations as in 1.7.

Since Prony's time this idea has been rediscovered several times in different fields. This technique is called the 'covariance method' of linear prediction in speech signal processing literature [39].

Since Prony's idea was intended for extracting signal parameters of a noiseless signal sequence y(n) (with w(n)=0 in formula 1.1), any observation noise tends to perturb the estimates of the parameters $\{s_k\}$ considerably [32]. If more than 2M data samples are available for processing, the effect of noise can be countered to some extent by using a 'least squares' approach, as can be expected. That is, the additional samples can be used to write more equations in 1.6, and the unknown $\{g_k\}$ can be found by satisfying these equations in the least square sense. This was attempted as early as 1924 in ref. [33]. See also Hildebrand [32] and Lanczos [78]. There have been numerous other modifications to Prony's method in which attempts have been made to counter the effects of noise. Van Blaricum [34,35] has compiled a long list of references relating to Prony's method. Prony's method and its relation to representation of signals by a sum of exponentials has been given considerable attention in refs. [36-38].

Many recently popularized parametric spectrum analysis techniques such as linear prediction analysis [39] and AR/ARMA modelling algorithms [3] bear similarities to Prony's method. However, these algorithms are concerned, in general, with obtaining a spectral estimate of the data sequence, instead of extracting specific signal parameters.

The indirect approach to estimating the signal parameters is attractive in comparison to other approaches due to at least three reasons. Firstly, it tends to be computationally less expensive. It involves solving a special set of linear equations for which computationally efficient algorithms are available [40,41], and rooting a polynomial G(z). In contrast, the direct approach detailed in section 1.3.1 involves a multidimensional search and the iterative

approach is not guaranteed to converge. Secondly, even if the signal underlying the data is not exactly a sum of exponentials as assumed, it is possible to draw meaningful inferences about the signal from the polynomial G(z) by evaluating it on the unit circle $|z|=1$. Thirdly, the resolution of closely spaced s_k's (which could represent the 'poles of a linear system' which generated the signal or frequencies of sinewaves) is limited primarily only by the SNR in the data. In other words, if the SNR is high enough, then arbitrarily closely spaced s_k's can be resolved or accurately determined. This is in stark contrast with the discrete Fourier transform based methods [42,19] in which the frequencies closer than 1/N cycles apart cannot be resolved even if the SNR is infinite. The situation is even worse if the signals are exponentially damped.

1.4: MAIN RESULTS:

Here we pursue the indirect approach originating from the work of Prony. That is, we compute the coefficients of a prediction-error filter polynomial G(z) from the N given samples of data and obtain the estimates of $\{s_k\}$ from M of its roots. The $\{a_k\}$ are subsequently estimated as in Prony's method. The parameter estimates obtained from Prony's method and its modifications are unsatisfactory in accuracy at low SNR. The term 'low SNR' is used here in a relative sense.

In this dissertation, we show how to obtain accurate estimates of $\{s_k\}$ (and subsequently $\{a_k\}$) from the zeros of the polynomial G(z) at low SNR values. We also give procedures for finding an estimate of M.

The improvement in parameter estimation accuracy achieved by our methods is primarily due to two reasons. They are briefly outlined below in the order of importance.

(1) We trade redundancy in the degree of G(z), L, for increased tolerance of parameter estimates to noise in the data. Ideally, if the data is noiseless,

an M th degree polynomial G(z) can be used to find $\{s_k\}$ from its M zeros as in Prony's method. But if the data is noisy, $\{s_k\}$ cannot be estimated accurately using an M th degree polynomial G(z). A value of L greater than M is necessary. This amounts to attempting to 'explain' the M signal components and the noise with L exponentials. This redundancy in the degrees of freedom tends to increase the accuracy in the parameter estimates. Several authors have alluded to this fact [43-49,18]. But this alone is not sufficient. To extract accurate signal parameter estimates, one has to distinguish which M of the L zeros of G(z) are related to the signal parameters (these M zeros are called signal zeros) and which L-M zeros (called extraneous zeros) are related to noise in the data. In chapter 3 we present a method for this purpose. We pick the M (signal) zeros out of the L zeros of G(z) using a simple least squares criterion. Using this procedure, Prony's method can be used at low SNR as well. This modification also provides a way of estimating the value of M.

(2) Further improvement in estimation accuracy is achieved by adaptive prefiltering using singular value decomposition (SVD) or eigenvalue-eigenvector decomposition. This part of the work is a summary of our published work [52-56]. This improvement stems from the following observation. The exact values of the exponent parameters $\{s_k\}$ can be obtained from the zeros G(z) if the data is noiseless, using Prony's method. Therefore it is reasonable to expect that if the SNR in the data can be improved by 'filtering', then more accurate parameter estimates can be obtained from the zeros of G(z). But to increase the SNR in the data some knowledge of the signal is needed. Although we do not know the value of the signal parameters, the fact that they are a sum of exponentials can be used for this purpose. Let the data samples y(n),n=1,2,...N be arranged in the form of a Toeplitz matrix A.

$$A = \begin{bmatrix} y(k) & y(k-1) & \cdots & y(1) \\ y(k+1) & y(k) & \cdots & y(2) \\ \vdots & \vdots & & \vdots \\ y(N-1) & y(N-2) & \cdots & y(N-k) \end{bmatrix} \qquad 1.8$$

If the data is noiseless, the rank of A will be M, provided that k satisfies certain constraints (see chapter 2). This property is peculiar to the exponential signals we are dealing with. This knowledge regarding the signal can be used to improve the SNR in the data. SVD provides a way of achieving this. Details are given in chapter 4.

1.5: ORGANIZATION OF THE DISSERTATION:

Chapter 2 is concerned with properties of the zeros of the prediction-error filter polynomial G(z) when its coefficients are computed from a noiseless (signal only) data sequence y(n). This chapter does not address the main problem of estimating the signal parameters $\{s_k\}$, $\{a_k\}$ and M. But the results obtained regarding the zeros of G(z) help provide insight into this problem.

Chapters 3 and 4 are concerned with algorithms for estimating the desired signal parameters from a noisy data sequence y(n). In chapter 3 we present a modification to Prony's method. This modification makes it possible to estimate the number of signal components M in the data and to accurately estimate the parameters at lower SNR values than previously possible with Prony's method.

In chapter 4 we present two methods based on singular value decomposition (SVD) to obtain even more accurate estimates. These methods are closely connected to several other techniques based on eigenvalue/eigenvector decomposition of an estimated autocorrelation matrix of data samples. An important contribution in chapter 4 is to describe these connections explicitly.

In chapter 5, computer simulation results are presented. The estimation accuracy, measured in terms of bias and variance of the parameter estimates of different methods, is calculated based on several independent trials and compared to the appropriate Cramer-Rao bounds.

In chapter 6, major conclusions are listed.

CHAPTER 2: ZEROS OF THE PREDICTION-ERROR POLYNOMIAL

Throughout this dissertation we have used the zero locations of a polynomial called the prediction-error filter polynomial G(z) to estimate the unknown exponent parameters s_k,k=1,2,...M. See section 1.1 for definition of the problem. The methods we have suggested differ primarily only in the manner in which the coefficients of G(z) are computed from the data samples y(1),y(2), ...y(N). Ideally, an M th degree polynomial G(z) should be sufficient to estimate the unknown parameters s_k,k=1,2,...M. However, when the data is noisy, as mentioned before, some redundancy in the degree, L>M, of G(z), improves the accuracy of the estimates of s_k [43-49,18]. In this chapter we shall study the properties of the zeros of G(z), as a function of different values of L, for fixed value of N, the number of data samples, and M, the number of signal components in the data. For analytical tractability we assume that the sequence y(n) is noiseless, i.e. that it is a sum of samples of M unknown exponential signals only. Although this assumption is unrealistic, the insights we gain by studying the properties of zeros of G(z) are helpful in the more practical case of noisy data, discussed in subsequent chapters.

There are two main points made in this chapter. Firstly, it is pointed out that, for a given value of N, there are upper and lower limits on the value of L, the degree of the polynomial G(z). What do we mean by these limits? If and only if L lies in between these limits can we hope to find the parameters s_k,k=1,2,...M from the zeros of G(z). This fact is important to keep in mind because in chapters 3 and 4, we try to improve the accuracy of the estimates of the parameters when the data is noisy, partly by using L values larger than M, and the limits on L derived in this chapter serve as bounds in that case. The

limits on L values are derived in section 2.1. Secondly, if L is chosen greater than M, the polynomial G(z) has L-M extra zeros (other than the M signal zeros which fall at e^{s_k}, k=1,2,...M). These L-M extraneous zeros turn out to be a necessary evil in our estimation algorithms presented in chapter 4. It is important to be able to identify the M signal zeros of G(z) from the L-M extraneous ones, so that the (estimates of) parameters s_k,k=1,2,...M can be found from the locations of the signal zeros. This is facilitated by imposing certain constraints on the coefficients of the polynomial G(z) as explained in sections 2.2 and 2.3. The effect of such constraints on G(z) is to cause the L-M extraneous zeros of G(z) to always fall inside the unit circle, $|z|=1$, in the z-plane. That is, these L-M zeros will have magnitude less than unity. This property is used in distinguishing the signal zeros from the extraneous ones. The constraints occur in a natural way when our singular value decomposition (SVD) based algorithms are used (See Chapter 4).

2.1: THE LOCATIONS OF THE SIGNAL ZEROS OF G(z):

In this section we shall show that the degree L of G(z) has maximum and minimum limits. The noiseless signal sequence y(n) is given by the formula

$$y(n) = \sum_{k=1}^{M} a_k e^{s_k n} \qquad n=1,2...N \quad 2.1$$

The exponent parameters s_k are assumed distinct. Let $\underline{g}'$ be a (L+1)x(1) vector of coefficients of a polynomial G(z).

$$\underline{g}' = (g_0, g_1, g_2 \cdots g_L)^T \qquad 2.2$$

$$G(z) = \sum_{k=0}^{L} g_k z^{-k} \qquad 2.3$$

'T' denotes matrix transpose. Let us arrange the data samples y(n) in the form of a (N-L)X(L+1) Toeplitz matrix A'.

$$A_f' = \begin{bmatrix} y(L+1) & y(L) & \cdots\cdots & y(1) \\ y(L+2) & y(L+1) & \cdots\cdots & y(2) \\ \cdot & \cdot & \cdots\cdots & \cdot \\ \cdot & \cdot & \cdots\cdots & \cdot \\ y(N) & y(N-1) & \cdots\cdots & y(N-L) \end{bmatrix} \qquad 2.4$$

Proposition 1:

If the coefficient vector $\underline{g}'$ satisfies $A_f' \underline{g}'=\underline{0}$, and if L satisfies the inequality $M\leq L\leq(N-M)$, then G(z) has M of its L zeros at e^{s_k}, k=1,2,...M.

This property is easy to see from the following observation. Any row of A_f' can be written as a linear combination of the following M linearly independent vectors.

$$\underline{f}_k = (1, e^{-s_k}, e^{-2s_k}, \ldots\ldots\ldots e^{-Ls_k}), \quad k=1,2,\ldots M \qquad 2.5$$

That is, the rank of A_f' is M, as long as A_f' has at least M rows, i.e., if $(N-L)\geq M$ or $L\leq(N-M)$. The null space of A_f' has dimension L+1-M. Since $\underline{g}'$ lies in the null space of A_f', $\underline{f}_k\underline{g}'=\underline{0}$, k=1,2,...M, if $L\leq(N-M)$. Therefore,

$$g_0+g_1e^{-s_k}+g_2e^{-2s_k}+\cdots\cdots+g_Le^{-Ls_k}=0, \quad k=1,2,\ldots M \qquad 2.6$$

In other words, the polynomial $G(z)=\sum_{k=0}^{L} g_k z^{-*}$ has zeros at e^{s_k}, k=1,2,...M.

However L has to be equal to or greater than M, in order that the null space of A_f' has at least dimension one. If L=M, the null space of A_f' has a unique vector which is proportional to that given by Prony's method [32]. If $L\geq(N-M)$ A_f' will have less than M rows (the rank of $A_f'<M$) and the property in equation 2.6 will not be true in general. Hence L should satisfy the inequality $M\leq L\leq(N-M)$ in order that G(z) may have zeros at e^{s_k}, k=1,2,...M.

Now let us define a (N-L)x(L+1) Hankel matrix, called a backward data matrix A_b', in which the N data samples y(n) are complex conjugated and arranged in reversed order compared to the forward data matrix defined in formula 2.4 '*' denotes complex conjugate.

$$A_b' = \begin{bmatrix} y^*(1) & y^*(2) & \cdots\cdots & y^*(L+1) \\ y^*(2) & y^*(3) & \cdots\cdots & y^*(L+2) \\ \cdot & \cdot & \cdots\cdots & \cdot \\ \cdot & \cdot & \cdots\cdots & \cdot \\ y^*(N-L) & y^*(N-L+1) & \cdots\cdots & y^*(N) \end{bmatrix} \qquad 2.7$$

Proposition 2:

If a coefficient vector $\underline{g}'$ satisfies the homogeneous equation $A_b'\underline{g}'=\underline{0}$, and if L satisfies the inequality $M \leq L \leq (N-M)$, the polynomial G(z) has zeros at $e^{-s_k^*}$, k=1,2,...M.

Following proposition 1, any row of A_b' can be written as a linear combination of the following M linearly independent vectors.

$$\underline{b}_k = (1,\ e^{s_k^*},\ e^{2s_k^*},\ \cdots\cdots\ e^{Ls_k^*}),\ k=1,2,\ldots M \qquad 2.8$$

if $L \leq (N-M)$. Thus, as before, $\underline{b}\ \underline{g}'=0$, k=1,2,...M. Therefore,

$$g_0+g_1e^{s_k^*}+g_2e^{2s_k^*}+\cdots\cdots+g_Le^{Ls_k^*}=0,\quad k=1,2,\ldots M \qquad 2.9$$

which means the polynomial $G(z)=\sum_{k=0}^{L} g_k z^{-k}$ has zeros at $e^{-s_k^*}$, k=1,2,...M. The rest of the argument is the same as in proposition 1. It is noted that these M signal zeros are the complex conjugates of the reciprocals of the corresponding M signal zeros in proposition 1. In other words, the M signal zeros of proposition 1 and 2 are the reflections of one another about the unit circle.

If the signal sequence y(n) is composed of undamped sinusoidal signals (i.e., all s_k's are purely imaginary numbers) then the polynomials in proposition 1 and 2 will have the same signal zeros, since e^{s_k} and $e^{-s_k^*}$. In this case, the rows of A_f' and A_b' are spanned by the same vectors, since $\underline{f}_k = \underline{b}_k$, k-1,2,...M. Thus the two data matrices can be combined to form the (2(N-L) x (L+1)) forward-backward data matrix, A_{fb}'.

$$A_{fb}' = \begin{bmatrix} A_f' \\ \hline A_b' \end{bmatrix} \qquad 2.10$$

Proposition 3:

If the vector $\underline{g}'$ satisfies $A'_{fb}\underline{g}'=\underline{0}$ and if L satisfies the inequality $M \leq L \leq (N- M/2)$, then the polynomial G(z) has M of its zeros on the unit circle at $e^{s_k}=e^{-s_k^*}$, k=1,2,...M. The data sequence y(n) is assumed to be composed of undamped sinusoids. M/2 stands for an integer rounded to the next larger integer value if M/2 is not an integer.

The arguments in this case are identical to those in proposition 1 or 2. The number of rows in A'_{fb} is 2(N-L). A'_{fb} should have at least M rows or $2(N-L) \geq M$ if G(z) is to have the property given in equation 2.6 or 2.9. Hence, L should satisfy $M \leq L \leq (N- M/2)$. Henderson's [57] result is closely related to our proposition 1. However, our results in this chapter differ from his in that we assume only a single data record of N samples is available for processing. The idea of using the data in both forward and backward directions has previously been used by Burg [58], Ulrych and Clayton [45], Nuttall [59], and others.

2.2: MINIMUM PHASE PROPERTY OF THE EXTRANEOUS ZEROS:

In the previous section we established that the polynomial G(z) has zeros at desired locations in the Z plane from which the signal parameters $\{s_k\}$ can be found if the degree of the polynomial satisfies certain inequalities. The polynomial G(z) also has L-M other zeros called extraneous zeros which turn out to be necessary in our estimation procedure. They are necessary because a polynomial G(z) with L>M gives more accurate estimates of the parameters $\{s_k\}$ if the data is noisy. But, they pose a problem since we have to identify the M signal zeros from the L-M extraneous zeros before estimating the parameters $\{s_k\}$. In this section we address the problem of finding a coefficient vector $\underline{g}'$ for the polynomial G(z) which would simplify this identification problem.

The location of the L-M extraneous zeros of G(z) depends on the choice of the $\underline{g}'$ vector which lies in the null space of A'_f (or A'_b or A'_{fb} as the case may be).

Recall that the dimension of the null space of A_f' (or A_b' or A_{fb}') is L+1-M. Hence there are at least L+1-M different choices for $\underline{g}'$. Since $A_f' \underline{g}'=\underline{0}$, $\underline{g}$ will also lie in the null space of $A_f'+A_f'$ where '+' stands for matrix complex conjugate transpose. Therefore, $\underline{g}'$ can be chosen as an eigenvector corresponding to one of the L+1-M zero eigenvalues of $A_f'+A_f'$. This approach is related to the Pisarenko method [60] and we shall say more about this in chapter 4. The problem with these choices of $\underline{g}'$ is that the L-M extraneous zeros of G(z) can fall anywhere in the Z plane and may not be identifiable from the M signal zeros in the absence of any prior information. The fact that the eigenvector has unit length, ($\|\underline{g}'\|$ =1) does not in any way constrain the location of the L-M extraneous zeros.

In the next proposition we find a unique vector $\underline{g}'$ (satisfying $A_f'\underline{g}'=\underline{0}$ or $A_f'\underline{g}'=\underline{0}$ or $A_{fb}'\underline{g}'=\underline{0}$ as the case may be) which results in a polynomial G(z) with certain desirable properties. In proposition 5 we show how to calculate such a vector $\underline{g}'$ from the given data samples.

Proposition 4:

The L-M extraneous zeros of the L th degree polynomial $G(z)=\sum_{k=0}^{L} g_k z^{-k}$ fall inside the unit circle, $|z|=1$, regardless of the M signal zero locations if its coefficient vector $\underline{g}'=(g_0, g_1, g_2, \ldots\ldots g_L)^T$ which satisfies $A_f'\underline{g}'=\underline{0}$ (or $A_b'\underline{g}'=\underline{0}$ or $A_{fb}'\underline{g}'=\underline{0}$) is chosen to have the following constraints: $g_0 = 1$ and $|g_1|^2+|g_2|^2+|g_3|^2+ \ldots\ldots +|g_L|^2$ is minimum. It is assumed that L satisfies the inequalities specified in proposition 1 (or 2 or 3 as the case may be).

As long as $\underline{g}'$ satisfies the homogeneous equation it follows from proposition 1 (or 2 or 3) that G(z) has M signal zeros at e^{s_k}, k=1,2,...M (or $e^{-s_k^*}$, k-1,2,...M). We can factor the polynomial G(z) into

$$G(z) = G_1(z)\, G_2(z) \qquad 2.11$$

where

$$G_1(z) = 1 + b_1 z^{-1} + b_2 z^{-2} + \ldots\ldots + b_M z^{-M}$$

$$G_2(z) = 1 + c_1 z^{-1} + c_2 z^{-2} + \ldots\ldots + c_{L-M} z^{-(L-M)}$$

The factor $G_1(z)$ is supposed to have the M signal zeros which fall at known locations as per propositions 1 or 2 or 3. We are however, interested in the locations of the extraneous zeros which are the zeros of the polynomial $G_2(z)$, when the coefficients of G(z) are chosen such that Q is minimum.

$$Q = 1 + |g|^2 + |g|^2 + \ldots\ldots + |g_L|^2 \qquad 2.12$$

Note that the first coefficient is chosen to be unity.

Since the polynomial G(z) is the product of $G_1(z)$ and $G_2(z)$, its coefficients are the convolution of the coefficients of $G_1(z)$ and $G_2(z)$. That is,

$$g_n = \sum_{k=-\infty}^{\infty} c_{n-k} b_k \qquad 2.13$$

where $g_0 = 1$, $b_0 = 1$, $c_0 = 1$, $b_i = 0$ for $i>M$ and $i<0$, and $c_i=0$ for $i>L-M$ and $i<0$. One can now imagine the above convolution, as performing linear prediction (This method of linear prediction is called the 'autocorrelation' method. See for example, ref. [39].) on the 'data sequence' $1, b_1, b_2, \ldots\ldots b_M$ (which are the coefficients of $G_1(z)$) using a 'prediction-error filter' defined by the vector of coefficients of the polynomial $G_2(z)$, $\underline{c}$, $\underline{c}=(1, c_1, c_2, \ldots\ldots c_{L-M})^T$. Then the coefficients of the polynomial G(z), $1, g_1, g_2, \ldots\ldots g_L$ form the 'error sequence' out of the filter $\underline{c}$. Thus the problem of minimizing Q in formula 2.12 is the same as one of finding a 'prediction-error filter' $\underline{c}$ that minimizes the error energy at its output, operating on a 'data sequence' $1, b_1, b_2, \ldots\ldots b_L$. It is well known that this problem results in a set of Toeplitz system of normal equations (Yule-Walker type, see ref. [39] for example). These equations are as below.

$$\begin{bmatrix} R(0) & R(1) & \cdots\cdots & R(L-M-1) \\ R(-1) & R(0) & \cdots\cdots & R(L-M-2) \\ \cdot & \cdot & \cdots\cdots & \cdot \\ \cdot & \cdot & \cdots\cdots & \cdot \\ R(M+1-L) & \cdots\cdots & & R(0) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ c_{L-M} \end{bmatrix} = - \begin{bmatrix} R(-1) \\ R(-2) \\ \cdot \\ \cdot \\ R(M-L) \end{bmatrix} \qquad 2.14$$

The elements of the Toeplitz matrix and the vector on the right side are given by the formula

$$R(i) = \sum_{k=0}^{M} b_k b_{k+i} \qquad 2.15$$

where b_k are the coefficients of $G(z)$ ($b_0=1$). The $R(i)$ will be zero beyond lag M since the 'data sequence' itself has only M+1 samples (b_k's). The solution to the above linear set of equations minimizes Q and determines the polynomial $G_2(z)$. It is well known [39] that such a polynomial has all its zeros (the L-M extraneous zeros of $G(z)$) inside the unit circle. A proof of this fact is given in Appendix A.

The above discussion is valid for any two polynomials $G_1(z)$ (with known zeros) and $G_2(z)$ (the zeros of which are chosen). Hence the following statement: If an L th degree polynomial with leading coefficient unity has M known zeros and the remaining L-M zeros are chosen to minimize the sum of magnitude squares of its L+1 coefficients, then the L-M chosen zeros will have magnitude less than unity. This, of course, is a restatement of the autocorrelation method of linear prediction.

Proposition 5:

The coefficient vector $\underline{g}'$ of the polynomial $G(z)$, satisfying the constraints mentioned in proposition 4, can be found by simply solving a linear system of equations.

Since $\underline{g}'$ satisfies the homogeneous equation $A_f'\underline{g}'=\underline{0}$ (or $A_b'\underline{g}'=\underline{0}$ or $A_{fb}'\underline{g}'=\underline{0}$) and $g_0=1$, we can rewrite the homogeneous equation as follows.

$$A_f \underline{g} = -\underline{h}_f \qquad 2.16$$

where $\underline{g}^T = (g_1, g_2, \ldots\ldots g_L)$, $\underline{g}' = (1, \underline{g}^T)^T$, and A_f' is partitioned as $A_f' = (\underline{h}_f \; A_f)$. It is easy to see that the columns of A_f and the column vector $\underline{h}$ lie in the span of the M linearly independent vectors, $(1, e^{s_k}, e^{2s_k}, \ldots\ldots e^{(L-1)s_k})^T$, k=1,2,...M. The rank of A_f is also M. Although there are many vectors which satisfy equation 2.16, the one which has minimum Euclidean length (or minimum value for $|g_1|^2 + |g_2|^2 + \ldots\ldots + |g_L|^2$ can be found as follows:

$$\underline{g} = -A_f^{\#} \underline{h}_f \qquad 2.17$$

where $A_f^{\#}$ is the pseudoinverse of A_f [61]. This vector $\underline{g}$ also minimizes $\|\underline{g}'\|$ with the constraint that $g_0=1$.

Now consider the situation when we find two coefficient vectors $\underline{g}_f'$ and $\underline{g}_b'$ using the data matrices A_f' and A_b', respectively, on the same signal sequence defined in formula 2.1. Let $\underline{g}_f'$ and $\underline{g}_b'$ satisfy $A_f' \underline{g}_f' = \underline{0}$ and $A_b' \underline{g}_b' = \underline{0}$, respectively, and have the constraints specified in proposition 4. Let the corresponding polynomials be called $G^f(z)$ and $G^b(z)$.

Proposition 6:

The L th degree ($(N-M) \leq L \leq M$) polynomials $G^f(z)$ and $G^b(z)$ have the same L-M extraneous zeros. This result can be used in identifying the M signal zeros from the L-M extraneous zeros in the presence of noise in the data.

From propositions 1 and 2 it follows that the M signal zeros of $G^f(z)$ are the reflections of the M corresponding signal zeros of $G^b(z)$ about the unit circle. If e^{s_k}, k=1,2,...M are the signal zeros of $G^f(z)$, then $e^{-s_k^*}$, k=1,2,...M are the signal zeros of $G^b(z)$. Let the polynomial factors corresponding to the signal zeros of $G^f(z)$ and $G^b(z)$ be named $G_1^f(z)$ and $G_1^b(z)$ respectively. Due to this special property of the signal zeros, the coefficients of $G_1^f(z)$ are the reversed, complex conjugated, scaled version of the coefficients of $G_1^b(z)$. That is, if the coefficients of $G_1^f(z)$ are $1, b_1, b_2, \ldots\ldots b_M$ then the coefficients of

of $G_1^b(z)$ are $(b_M^*, b_{M-1}^*, \ldots\ldots b_1^*, 1)/b_M^*$. Now, recall that the location of the L-M extraneous zeros are determined by the solution of the equations in 2.14. The solution vector $\underline{c}$ depends on the values of R(i)'s. However, looking at the way in which the R(i)'s are computed in formula 2.15, it is clear that they are unaffected by the reversal and complex conjugation of b_k's. And the scale factor $1/b_M^*$ occurring in the coefficients of $G_1^b(z)$ affects both sides of 2.14 equally, leaving the vector $\underline{c}$ unchanged. Therefore we conclude that the L-M extraneous zeros of $G^f(z)$ and $G^b(z)$ remain the same while the signal zeros are reflected in radius about the unit circle as the prediction direction is changed (i.e., the homogeneous equations $A_f'\underline{g}'_f=\underline{0}$ and $A_b'g_b'=\underline{0}$ are used) on a given data sequence.

When s_k are purely imaginary (sinusoid case), e^{s_k} and $e^{-s_k^*}$ are the same. Hence the polynomials $G_1^f(z)$ and $G_1^b(z)$ (the factors of $G^f(z)$ and $G^b(z)$ respectively) are the same. Thus the polynomials $G^f(z)$, $G^b(z)$ (and also $G^{fb}(z)$) have the same signal and extraneous zeros and are, in fact, identical.

2.3: SPREAD OF THE EXTRANEOUS ZEROS:

In the previous section we showed that the L-M extraneous zeros of G(z) fall inside the unit circle if the polynomial coefficient vector $\underline{g}'$ obeys certain constraints. We shall now give a qualitative argument to show that these extraneous zeros are approximately uniformly distributed around the inside of the unit circle. We also give some examples to illustrate this point.

In equation 2.14 the R(i)'s are zero beyond lag M. Therefore, one can imagine the R(i) sequence to be a legitimate autocorrelation sequence of a moving average (MA) process. This hypothetical MA process would be the output of an all zero filter excited by white noise, the zeros of which coincide with the M signal zeros. Further, we can imagine that the 'prediction-error filter'

$\underline{c}$, is attempting to whiten this process, invoking the well-known whitening property of the autocorrelation method of linear prediction [39]. In fact, an infinite order filter $\underline{c}$ will be required to perfectly whiten an MA process. Since $\underline{c}$ attempts to whiten the spectrum (which is the Z transform of the sequence R(i) evaluated on the unit circle) it has to have zeros (the extraneous zeros of G(z)) around the unit circle wherever there is concentration of power, i.e., wherever the Z transform of the sequence R(i) does not have a (signal) zero close to the unit circle.

We now give an example in which the signal sequence is given by the formula $y(n) = e^{s_1 n} + e^{s_2 n}$, n=1,2,..... where $s_1 = (0.1+j2\Pi(0.35))$, $s_2 = (0.0+j2\Pi(0.5))$ and $s_3 = (-0.3+j2\Pi(0.65))$. The signal zeros of $G^f(z)$ (which can be found by solving $A'_f\underline{g}'_f=\underline{0}$) are located at e^{s_1}, e^{s_2}, e^{s_3}. The factor $G_1^f(z)$ of $G^f(z)$ is $\prod_{k=1}^{3} (1-e^{s_k}z^{-1})$. The plot of the corresponding 'MA spectrum' ($|G_1^f(z)|^2$, $z=e^{jw}$) is shown in figure 2.1. L was chosen 50. The factor $G_2^f(z)$ of $G^f(z)$ corresponding to the L=3 (=47) extraneous zeros was computed. It is interesting to note that this factor can be computed by solving the Toeplitz equations in 2.14 by using Levinson's recursions [39]. The magnitude of the reciprocal of $G_2^f(z)$ evaluated on the unit circle is plotted in figure 2.2 Figure 2.3 shows the 47 extraneous zeros and 3 signal zeros of $G^f(z)$. Note that the extraneous zeros are approximately uniformly distributed within the circle. However, if the data is noisy, the extraneous zeros tend to fall closer to or outside the unit circle and are usually observed as spurious spikes in spectral estimate plots (see for example ref. [49]).

If the data sequence is composed of exponentially damped signals, and if we use the forward data matrix A'_f (i.e., if we solve $A'_f\underline{g}'_f=\underline{0}$ to find the polynomial $G^f(z)$), we notice that both the signal zeros and the extraneous zeros

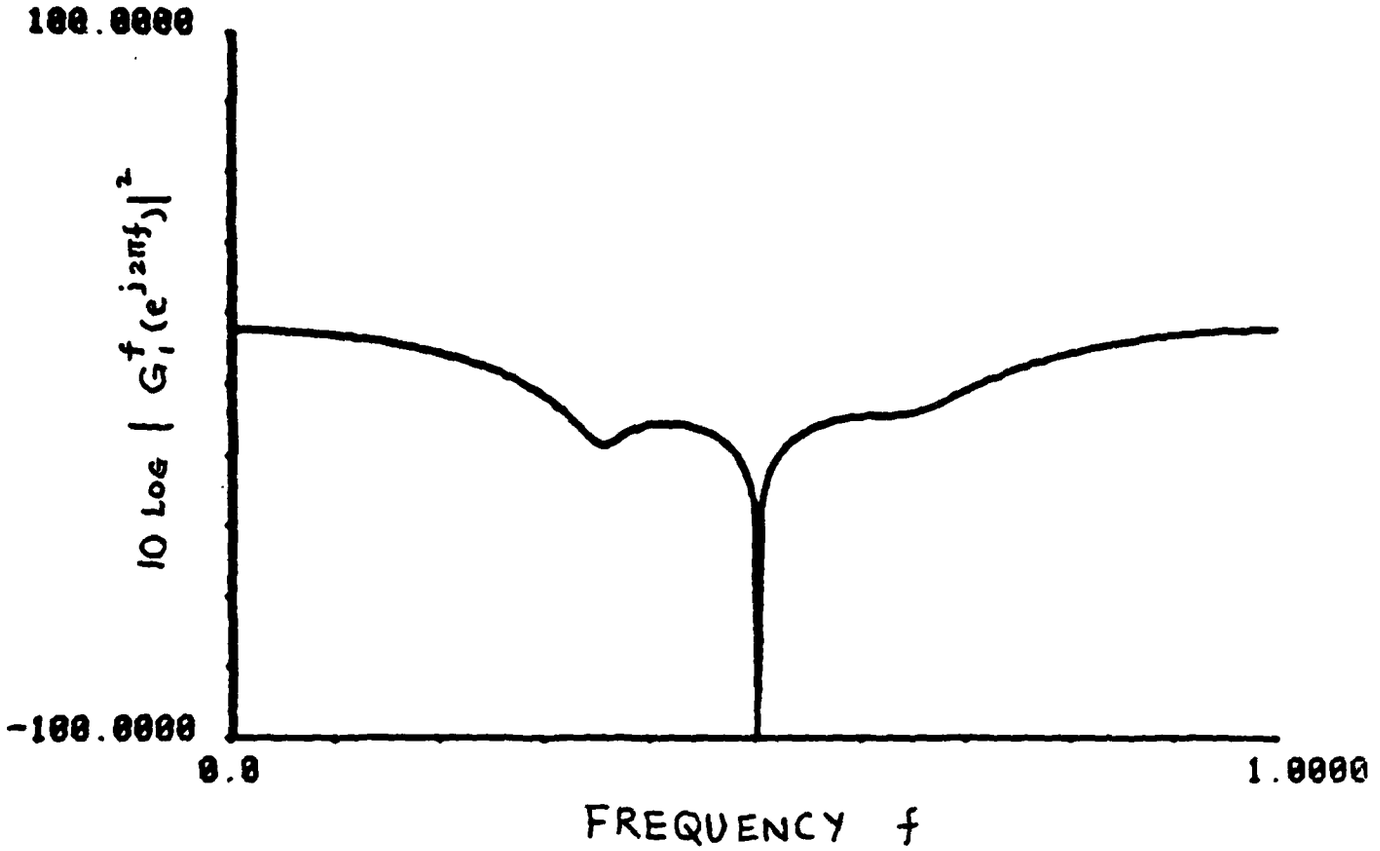


Figure 2.1: The magnitude of the polynomial $G_1^f(z)$ on the unit circle. It is proportional to the spectrum of the hypothetical MA process.

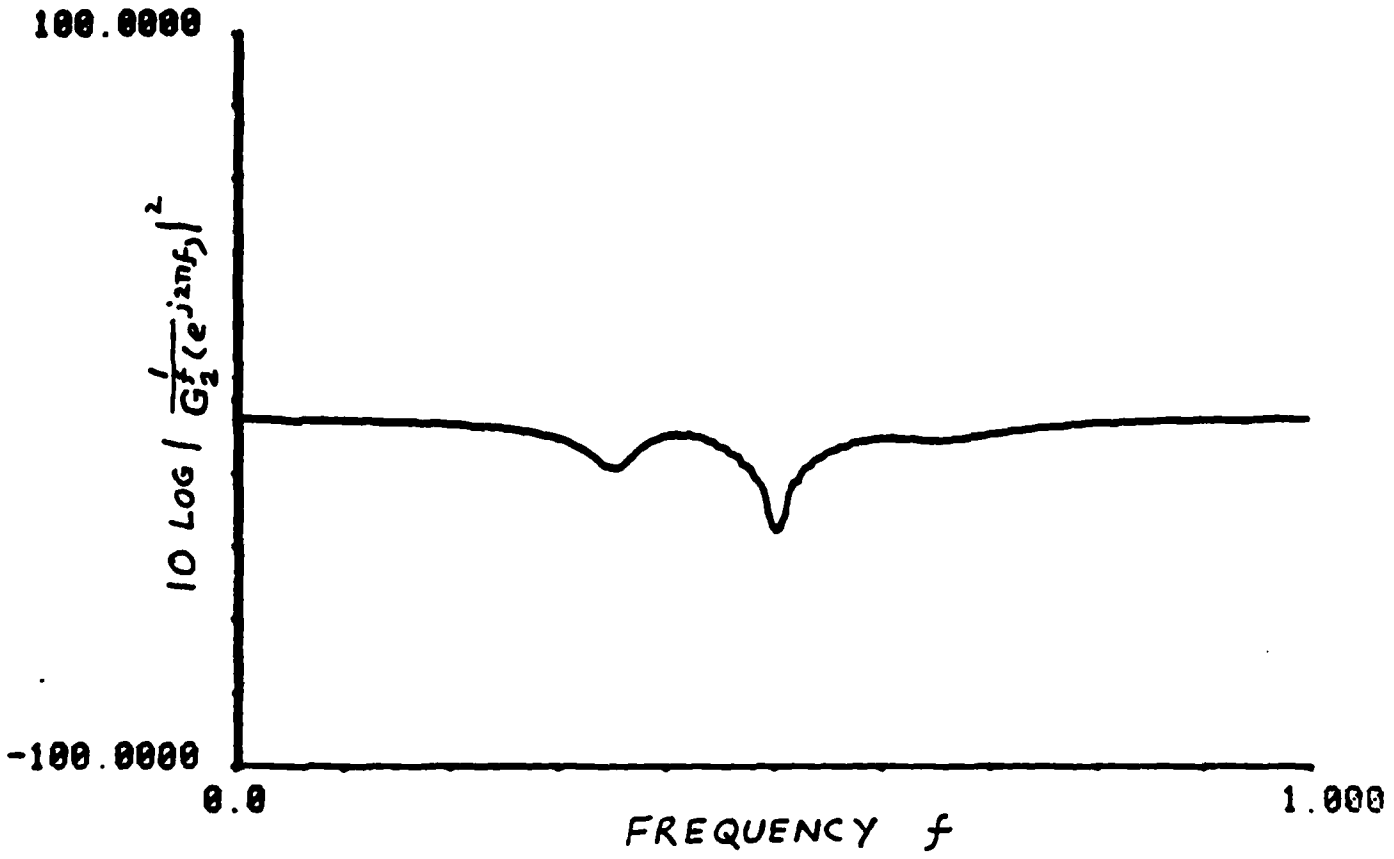


Figure 2.2: The magnitude of reciprocal of the 'prediction-error filter' polynomial $G_2^f(z)$. As L is increased $(1/|G_2^f(z)|^2)$ tends to resemble $|G_1^f(z)|^2$more closely.

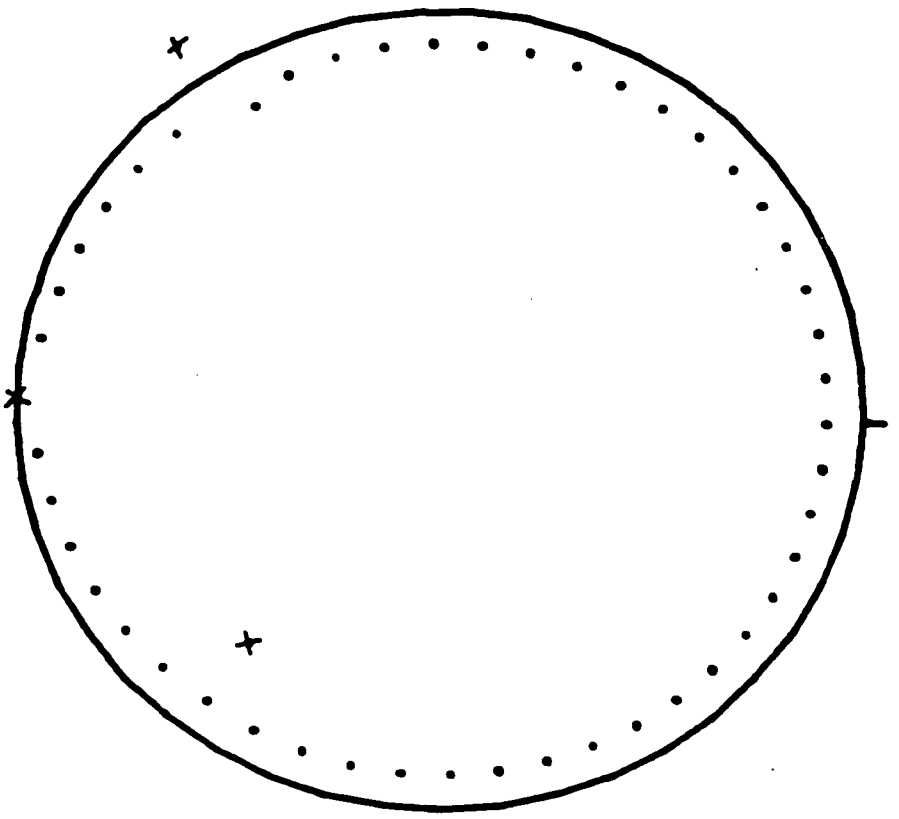


Figure 2.3: The signal zeros (denoted by x) and 47 extraneous zeros (denoted by dots) of G(z).

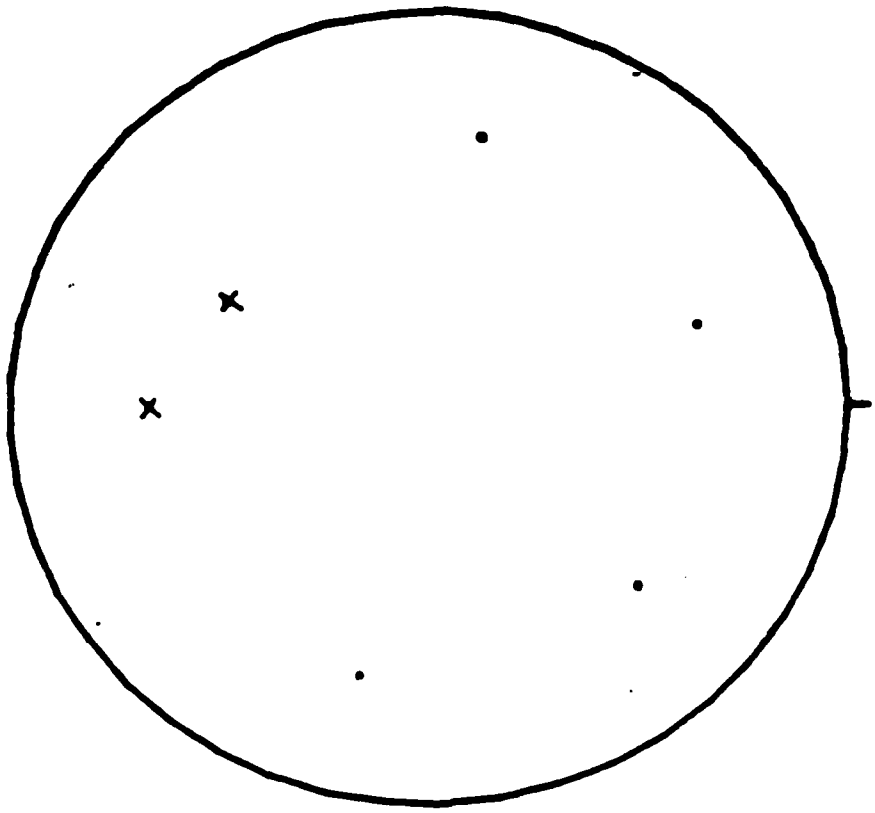


Figure 2.4: The zeros of $G^f(z)$ for a signal $y(n)=e^{(-0.4+j2\Pi(0.5))n}+e^{(-0.6+j2\Pi(0.42))n}$, $n=1,2,\ldots 8$. The equations $A_f\underline{g}=-\underline{h}_f$ were used. L=6.

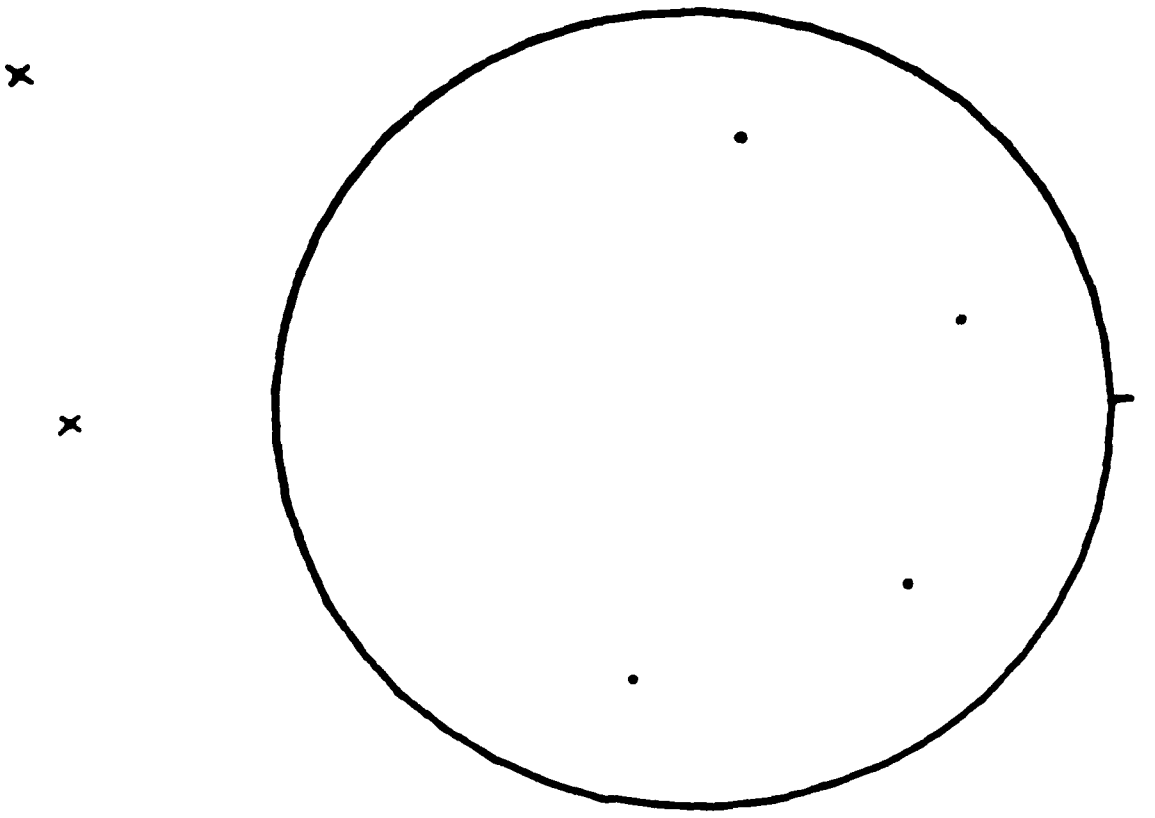


Figure 2.5: The zeros of $G^b(z)$. Same signal sequence as used in figure 2.4. The equation $A_b \underline{g} = -\underline{h}_b$ is used. L=6, Note that only signal zeros fall outside.

have magnitude less than unity. In this case the signal zeros cannot be identified without further information. Instead, using the backward data matrix A_b' (i.e., using the polynomial $G^b(z)$) one can distinguish the M signal zeros since only they fall outside the unit circle. This fact is made use of in chapters 4 and 5. An example is given in figures 2.4 and 2.5.

2.4: SUMMARY:

The degree L, of G(z) has to lie in between M and N-M (or N-M/2 in the case of forward-backward equations). Only then the signal parameters can be extracted from the M zero locations of G(z). If the coefficients of G(z) are chosen such that $g_0=1$ and $|g_1|^2 + |g_2|^2 + \ldots\ldots + |g_L|^2$ is minimum, then the L-M extraneous zeros always fall inside the unit circle. This fact can be used to identify which M of the L zeros of G(z) are signal zeros.

CHAPTER 3: MODIFICATIONS TO PRONY'S METHOD

In this and the next chapters we present actual algorithms for estimating the unknowns M, $\{s_k\}$, and $\{a_k\}$ from N samples of noisy data. In this chapter, we present a rather simple modification to Prony's method. The advantages of our modifications are 1) an estimate of the value of M can be found unlike in Prony's method; 2) the method can now be used at lower SNR values than previously possible. It is interesting to note that this particular modification to Prony's method had not previously been attempted, to our knowledge, in spite of the vast literature [34,35] on this topic.

Two key ideas in this modification are as follows. Firstly, using Prony's method, we try to 'fit' L (L>M) exponentials to the data sequence consisting of M exponentials and noise. This has been done by several authors [43-50,18] and is the primary reason for improved accuracy at low SNR. Subsequently we identify which M of the L exponentials correspond to the signal part of the data, using a least squares criterion.

We shall first describe Prony's method briefly. The modifications we propose are described in section 3.2.

3.1: PRONY'S METHOD AND SOME CLOSE RELATIVES:

Prony's method (according to Hildebrand [32]) attempts to 'fit' a model consisting of linear combination of exponentials to a data sequence. That is $\{a_k\}$ and $\{s_k\}$ are determined such that,

$$y(n) \approx \sum_{k=1}^{L} \hat{a}_k e^{\hat{s}_k n} \qquad n=1,2,\ldots N \qquad 3.1$$

$\{\hat{s}_k\}$ are obtained from L zeros of an L th degree polynomial G(z). The coefficients of $G(z)$ $(=1 + \sum_{k=1}^{L} g_k z^{-k})$, $g_1, g_2, \ldots g_L$ are determined by approximately satisfying the following set of equations.

$$\sum_{k=1}^{L} y(n-k)\, g_k = -y(n), \qquad n=L+1,L+2,\ldots N \qquad 3.2$$

We shall call these linear prediction equations since the (L+1)st sample is 'predicted' by a linear combination of the L samples and so on. The 'rationale' for attempting to satisfy 3.2 is that if 3.1 is satisfied with equality, 3.2 will also be satisfied with equality and exact values of $\hat{s}_k$'s can be found from the roots of G(z). In matrix notation of chapter 2 the above 'forward' linear prediction equations can be written as follows.

$$A_f \underline{g} \cong -\underline{h}_f \qquad 3.3$$

where

$$A_f = \begin{bmatrix} y(L) & y(L-1) & \ldots\ldots & y(1) \\ y(L+1) & y(L) & \ldots\ldots & y(2) \\ . & . & \ldots\ldots & . \\ . & . & \ldots\ldots & . \\ . & . & \ldots\ldots & . \\ y(N-1) & y(N-2) & \ldots\ldots & y(N-L) \end{bmatrix}, \quad \underline{g}^T=[g_1, g_2, \ldots g_L], \quad \underline{h}^T_f = [y(L+1), y(L+2), \ldots y(N)]$$

Prony's method (Hildebrand version) minimizes the equation error between the two vectors on either side of the equality sign in equation 3.3, which is $\|A_f\underline{g} + \underline{h}_f\|$ ($\|\underline{x}\|$ stands for the L_2 norm of $\underline{x}$.) by finding a coefficient vector $\underline{g}$. This error is also called prediction error and Prony's method is also called the 'covariance method' of linear prediction [39]. Such a vector $\underline{g}$ is found by using the pseudoinverse [61] of A_f, i.e.

$$\underline{g} = -[A_f^{\dagger}A_f]^{-1} A_f^{\dagger}\underline{h}_f \qquad 3.4$$

After computing the $\underline{g}$ vector, the polynomial $G(z) = 1 + \sum_{k=1}^{L} g_k z^{-k}$ is formed. Its zeros $e^{\hat{s}_k}$, k=1,2,...L are calculated. The $\hat{a}_k$'s are then found by satisfying equation 3.1 in the least square sense.

Generally, Prony's method will achieve an inferior 'fit' to the data compared to the direct least squares approach taken in sections 1.3.1 and 1.3.2

because Prony's method attempts to minimize an equation error $\|A_f \underline{g} + \underline{h}_f\|^2$ instead of the true least square error $\sum_{n=1}^{N} |y(n) - \sum_{k=1}^{L} \hat{a}_k e^{\hat{s}_k n}|^2$. In a special case, when the data is composed of purely exponential signals, (the minimum value of) both errors will be zero and both methods will achieve an exact 'fit'. Recall that in chapter 2 we considered this (noiseless) case in detail, where equation 3.3 was satisfied with equality.

In the above, instead of using the 'forward' prediction equations, it is possible to write down the 'backward' prediction equations as

$$\sum_{n=1}^{L} y^*(n+K)\, g_k \cong y^*(n), \quad n=1,2,\ldots N-L \qquad 3.5$$

or in matrix form (see chapter 2)

$$A_b \underline{g} \cong -\underline{h}_b \qquad 3.6$$

and achieve a different 'fit' to the data. The coefficient vector $\underline{g}$ obtained by (approximately) satisfying 3.6 ($\underline{g} = (A_b^{\dagger} A_b)^{-1} A_b^{\dagger} \underline{h}$) will be, in general, different from that obtained in 3.4. Going one step further, both forward and backward equation errors can be minimized simultaneously, by finding a common vector $\underline{g}$. That is, the equations $A_{fb}\underline{g} \cong -\underline{h}_{fb}$ (see chapter 2) can be used [45,59]. These modifications may or may not provide a better 'fit' to the data depending upon the data sequence y(n), n=1,2,...N. But they are helpful in signal parameter estimation in particular situations. In chapter 2, we mentioned that by using backward prediction equations one might be able to identify the M signal zeros of G(z) from others. Several authors [45,46,47,50] have found that using forward and backward linear prediction equations simultaneously gives more accurate parameter estimates, when the data is composed of undamped sinewaves in noise. These advantages are exploited in the sequel.

We pause a moment to point out that the coefficient vector $\underline{g}$ (in both cases of $A_f \underline{g} \cong -\underline{h}_f$ and $A_{fb}\underline{g} \cong -\underline{h}_{fb}$) can be computed using fast algorithms developed by Morf et al. [40] and Marple [41]. Now, we shall describe the specific modifications we propose and the need for them.

3.2: PROPOSED MODIFICATIONS:

The modifications we propose are in two steps. The first step is well known [43-50,18].

3.2.1: USING A VALUE FOR L GREATER THAN M:

If the data consists of M signal components and noise, it is unlikely that one can 'fit' L=M exponentials as in 3.1 to the data sequence using Prony's method and obtain accurate estimates of the signal parameters. This is because the equations 3.2 (with L=M) are satisfied only in the absence of noise in the data. Instead, one can 'fit' using Prony's method L>M exponentials to the noisy data sequence. Although the true value of M is not known usually in practical situations, such as in speech signal processing [2], an upper limit on M can be assumed and L can be chosen appropriately. If the number of data samples is relatively small (say, 30), L can be chosen approximately equal to N/2. Of course, L should lie within the upper limit of N-M (or N-M/2) as pointed out in chapter 2.

Intuitively, the L-M additional exponentials we 'fit' to the data tend to model the noise part of the data. This tends to improve the accuracy with which the M other exponentials in the 'fit' model the signal part of the data, eventually leading to more accurate parameter estimates. A problem which ensues from the use of L>M exponentials is the difficulty in determining which M of the L exponentials corresponds to the signal part of the data.

The improvement in accuracy achieved by using a value of L>M and the associated problems are best demonstrated by the following examples. In the first example, we simulate in a computer a sequence y(n), n=1,2,...25 consisting of two complex, undamped sinewaves in white, complex Gaussian noise. We use a modification of Prony's method, described above. That is, the equations

$A_{fb}\underline{g} = -\underline{h}_{fb}$ are solved to find $\underline{g}$ (i.e., $\underline{g} = -(A_{fb}^{\dagger}A_{fb}^{\dagger})^{-1} A_{fb}\underline{h}_{fb}$). The polynomial $G(z) = 1 + \sum_{k=1}^{L} g_k z^{-k}$ is formed and its zeros are found. These zeros, obtained over 50 trials in which different noise epochs were used, are plotted in figures 3.1-3.3 with respect to the unit circle $|z| = 1$. The true zero locations (on the unit circle) corresponding to the two complex sinewaves are shown by arrows. In figure 3.1 we chose L=2. That is G(z) is a second degree polynomial. If the data were noiseless, this would be adequate to find the frequencies of the sinewaves. But the noise in the data causes the zeros of G(z) to fall far away from their true locations. In figure 3.2, we chose L=12. We used the same data sequences as before. Clearly, there are two clusters of zeros around the true frequency locations, indicating improvement in estimation accuracy. The situation at lower SNR is shown in figure 3.3

Although providing redundancy in L improves the accuracy of the zero locations and hence the parameter estimates, the L-2 extraneous zeros cause a dilemma. At high SNR, these zeros behave as predicted in chapter 2 and are distributed within the unit circle. Hence, as usually done [47,49], one can find the zeros closest to the circle and use their angular location to estimate the frequencies. Alternatively, the location of the largest 'spectral peaks' in a 'spectral estimate', which is the magnitude of the reciprocal of G(z) evaluated on the unit circle, can be used as frequency estimates [45]. But at low SNR, some of the L-2 extraneous zeros of G(z) also fall close to the unit circle, and could cause large error in the frequency estimates. These zeros cause the large spurious spectral peaks observed in 'spectral estimate' plots when a large value of L is used (see for e.g. ref. [46]). The problem is even worse if the data is composed of exponentially damped sinusoids as seen in figures 3.4 and 3.5.

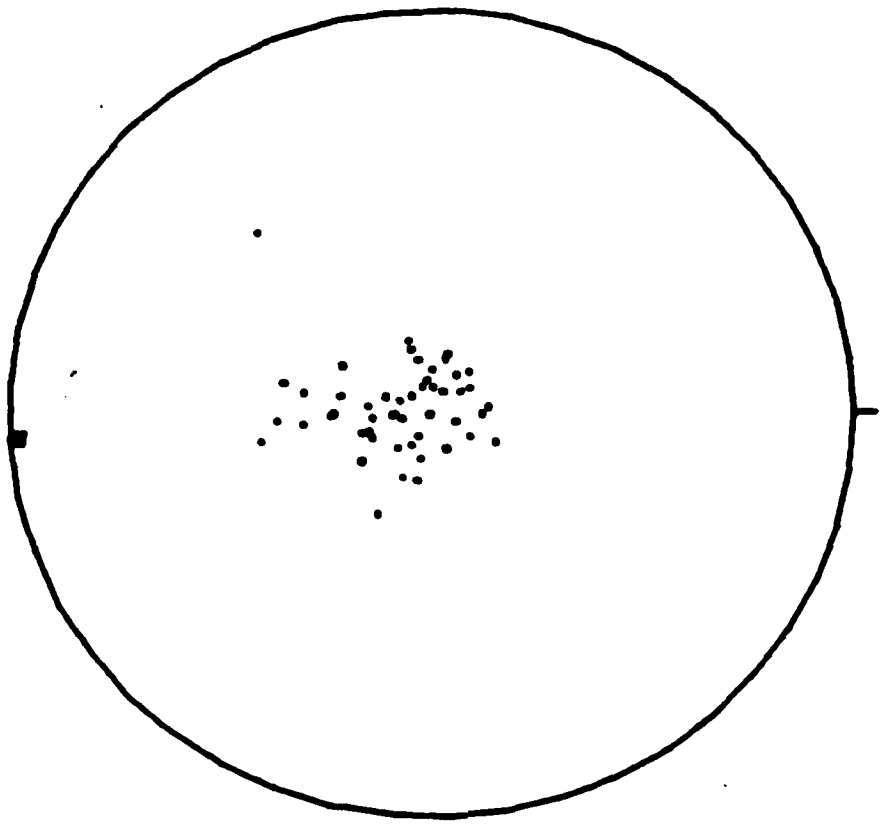


Figure 3.1: The zeros of G(z) for 50 independent trials. The data sequence $y(n+1)=e^{j2\Pi f_1 n+j\phi_1}+e^{j2\Pi f_2 n+j\phi_2}+w(n)$, $n=0,1,2,\ldots 24$, $f_1=0.52$Hz, $f_2=0.5$Hz, $\phi_1=\Pi/4$, $\phi_2=0$, SNR is defined as 10 Log $(1/2\sigma^2)$, where $2\sigma^2$ is the variance of the complex valued noise samples w(n). SNR=20 db. L was chosen 2.

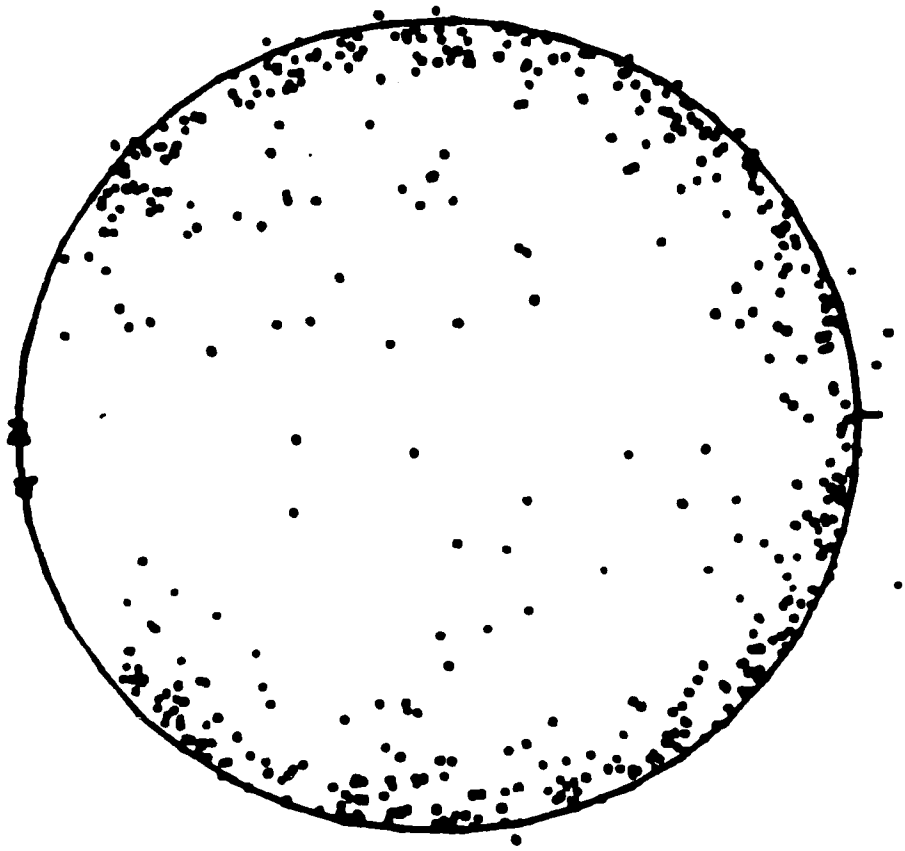


Figure 3.2: The zeros of G(z). The same data sets as in figure 3.1 are used. L is chosen 12. SNR=20 db.

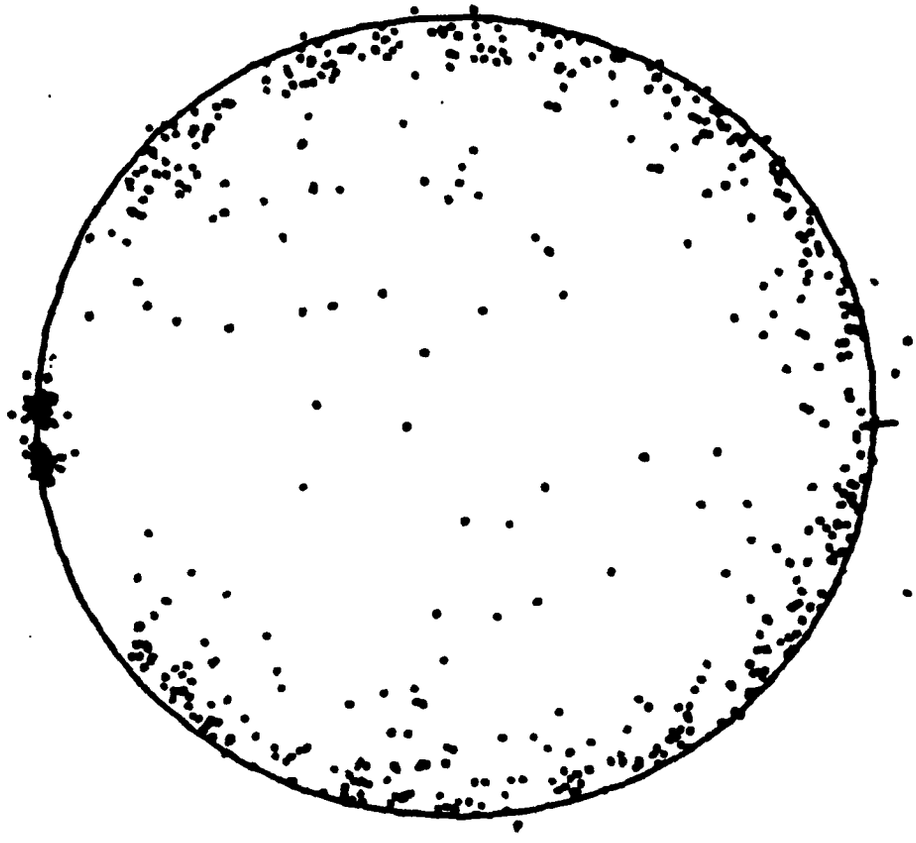


Figure 3.3: The zeros of G(z), L=12, SNR=10 db.

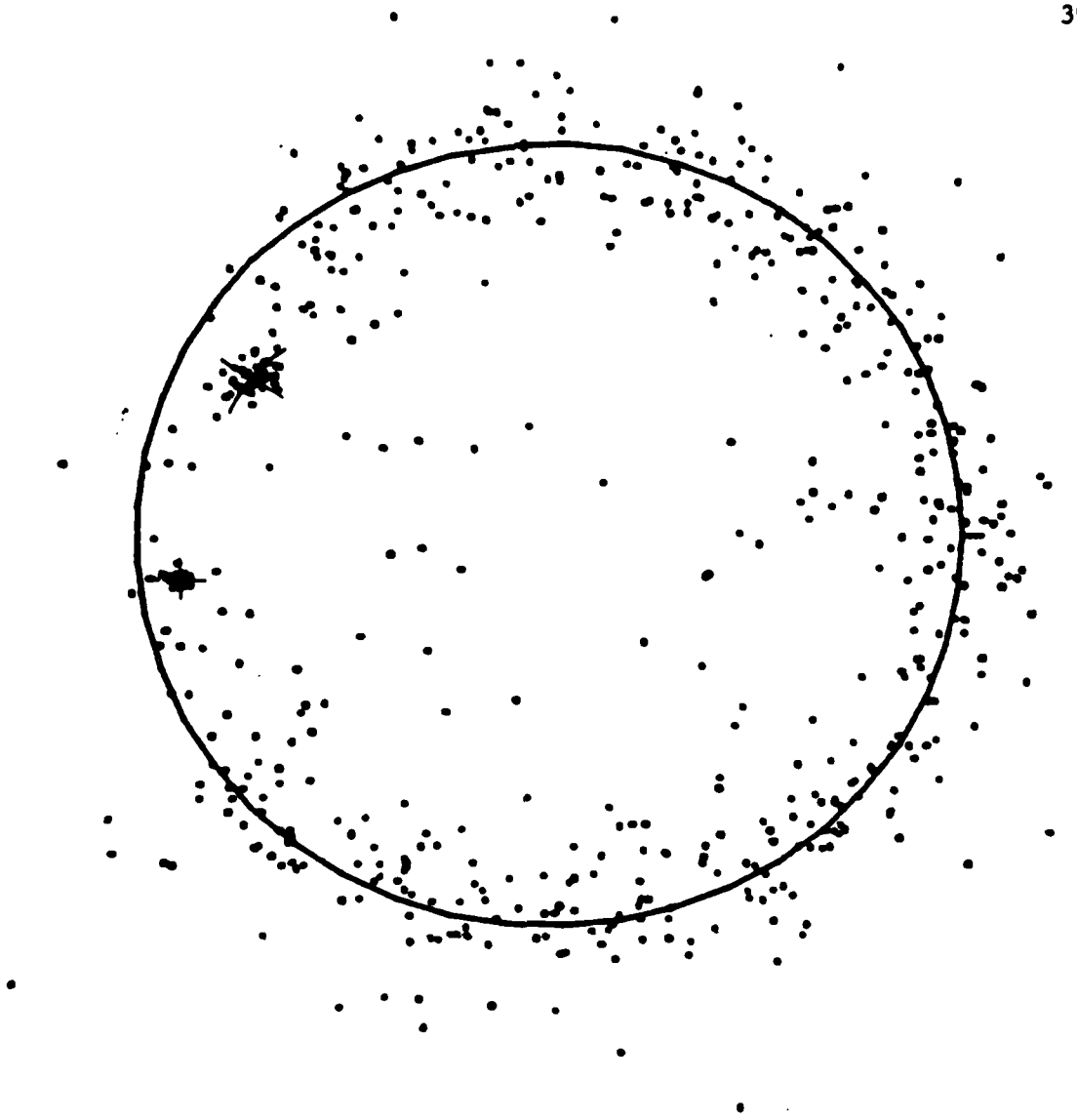


Figure 3.4: The zeros of G(z). The data sequence used: $y(n+1)=e^{s_1 n}+e^{s_2 n}+w(n)$, $n=0,1,2,\ldots 24$. SNR=20 db. SNR is defined as 10 log $(1/2\sigma^2)$ where $2\sigma^2$ is the variance of any noise sample w(n). The forward linear prediction equations are used, i.e., $\underline{g}=-(A_f^{\dagger}A_f)^{-1}A_f^{\dagger}\underline{h}_f$. Note the two signal zero clusters. The true locations are marked by 'x'., $s_1=-0.1+j2\Pi(0.52)$; $s_2=-0.2+j2\Pi(0.42)$.

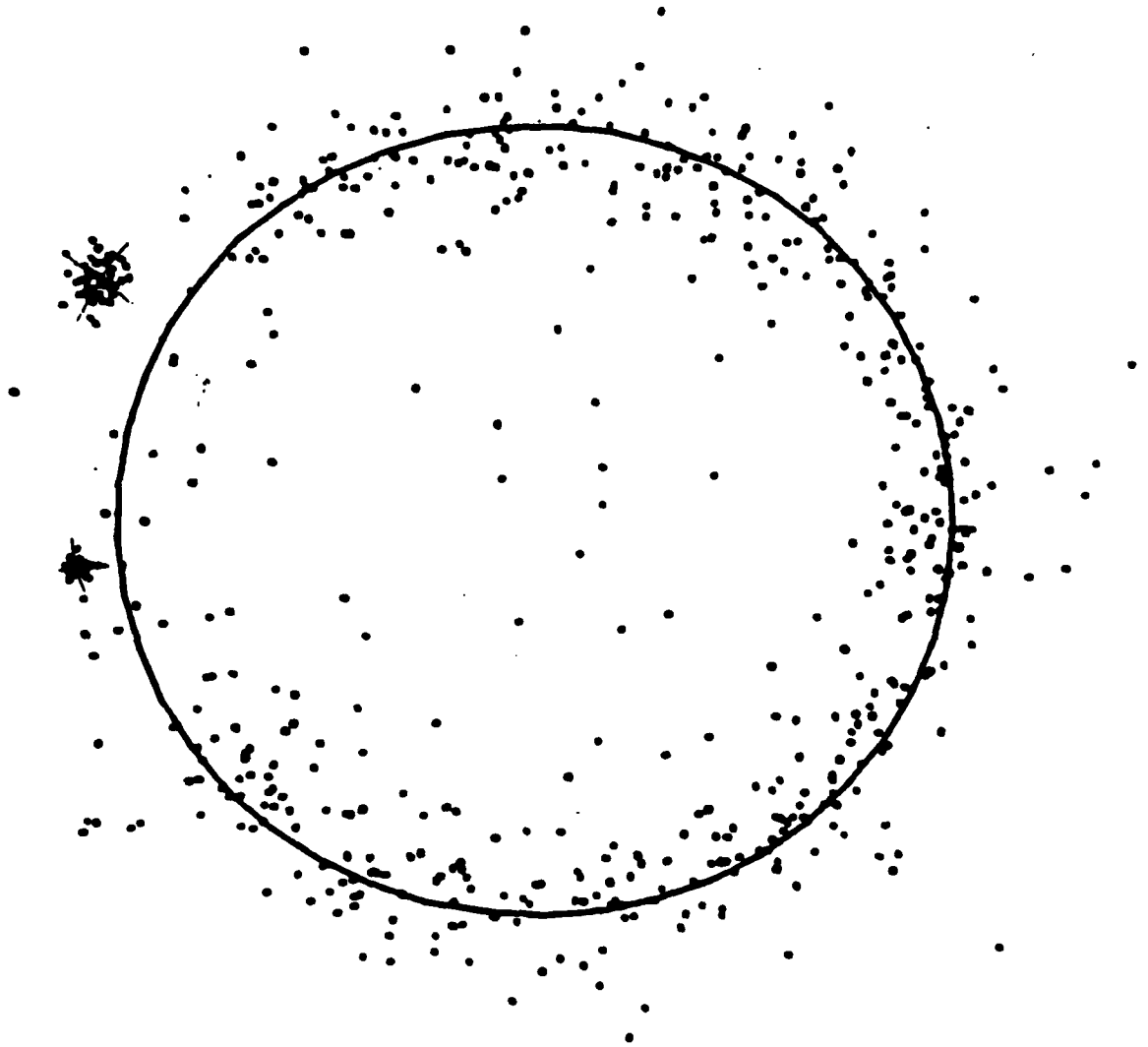


Figure 3.5: The zeros of G(z). The same data sets as in figure 3.4 are used. The backward linear prediction equations are used, i.e. $g=-(A_b^{\dagger}A_b)^{-1}A_b^{\dagger}h_b$.

In the next subsection we suggest a method to eliminate the L-M extraneous zeros using a least squares criterion.

3.2.2: ESTIMATING THE VALUE OF M AND ELIMINATING THE L-M EXTRANEOUS EXPONENTIALS:

Up to this point our procedure is essentially identical to Prony's method. Now we make use of our knowledge regarding the data sequence, i.e., that it contains M exponentials and relatively minor random errors. Secondly, the L zeros of G(z) which correspond to L exponentials $e^{\hat{s}_k}$, k=1,2,...L contain the desired exponentials which model the signal part of the data. Thus we seek to find M (if M is not known, its estimate $\hat{M}$ can be found as explained below) out of the L exponentials, $e^{\hat{s}_{k_i}}$, $1 \leq k_i \leq L$, i=1,2,...M that best fit the data, in the sense that they give minimum value for E, defined below.

$$E = \sum_{n=1}^{N} \left| y(n) = \sum_{i=1}^{M} \hat{a}_{k_i} e^{\hat{s}_{k_i} n} \right|^2 \quad \underset{a_{k_i},\ i=1,2,\ldots M}{\text{Min}} \qquad 3.7$$

The above minimization involves choosing combinations of M exponentials out of the L, solving a linear least square problem for the $\hat{a}_{k_i}$, and checking $\binom{L}{M}$ combinations to find the best combination, i.e. the one which gives the minimum value for E. The corresponding amplitude estimates are obtained as a by product. Some of the L exponentials can be eliminated to start with, using prior information. For example, if one is looking for frequencies of undamped sinewaves, those $e^{\hat{s}_k}$'s which have magnitude far greater than or far less than unity can be eliminated before choosing the best subset of size M. Also, an efficient subset selection procedure due to Hocking and Leslie [62] can be used to reduce computation. The author came to know about this procedure from the work of Tufts and Khorrami [63].

If M is not known apriori an estimate of M called $\hat{M}$, can be found as follows. Choose $\hat{M}$=1, and find the best subset of size unity that best fits the data. Call the corresponding minimum error E_1. Then, choose $\hat{M}$=2 and find the

best subset of size two and the corresponding minimum error E_2. Repeat the procedure until the rate of decrease of the error with increasing values of $\hat{M}$ is small, consistent with the modelling of the residual random errors. The integer i at which E_i shows the significant drop in rate of decrease is taken as $\hat{M}$. An example showing the value of the minimum error vs. $\hat{M}$ is given in figure 3.6. Simulation results using these modifications are presented in chapter 5.

3.3: DISCUSSION:

Ideally, to fit exponentials to a data sequence y(n) one has to minimize the error $\sum_{n=1}^{N} |y(n) - \sum_{k=1}^{M} \hat{a}_k e^{\hat{s}_k n^2}|$ with respect to $\hat{a}_k$'s and $\hat{s}_k$'s simultaneously. This is a difficult problem, even if the value of M is known. Instead, we find the exponentials $\hat{e}^{s_k}$'s separately as is often done, using Prony's method. However, we have made use of the fact that if the data is composed of exponentials and noise, overestimating the degree L (>M) of the polynomial G(z) improves the accuracy of the M signal zero locations. Subsequently, we select the M out of L exponentials (or zeros) that best explain the data. Of course, this step assumes that the random errors w(n) in formula 1.1 are relatively minor compared to the signal energy. But this assumption has to be valid to a lesser or greater degree for any estimation procedure to work. Using these modifications Prony's method can be used at (relatively) lower SNR than previously possible.

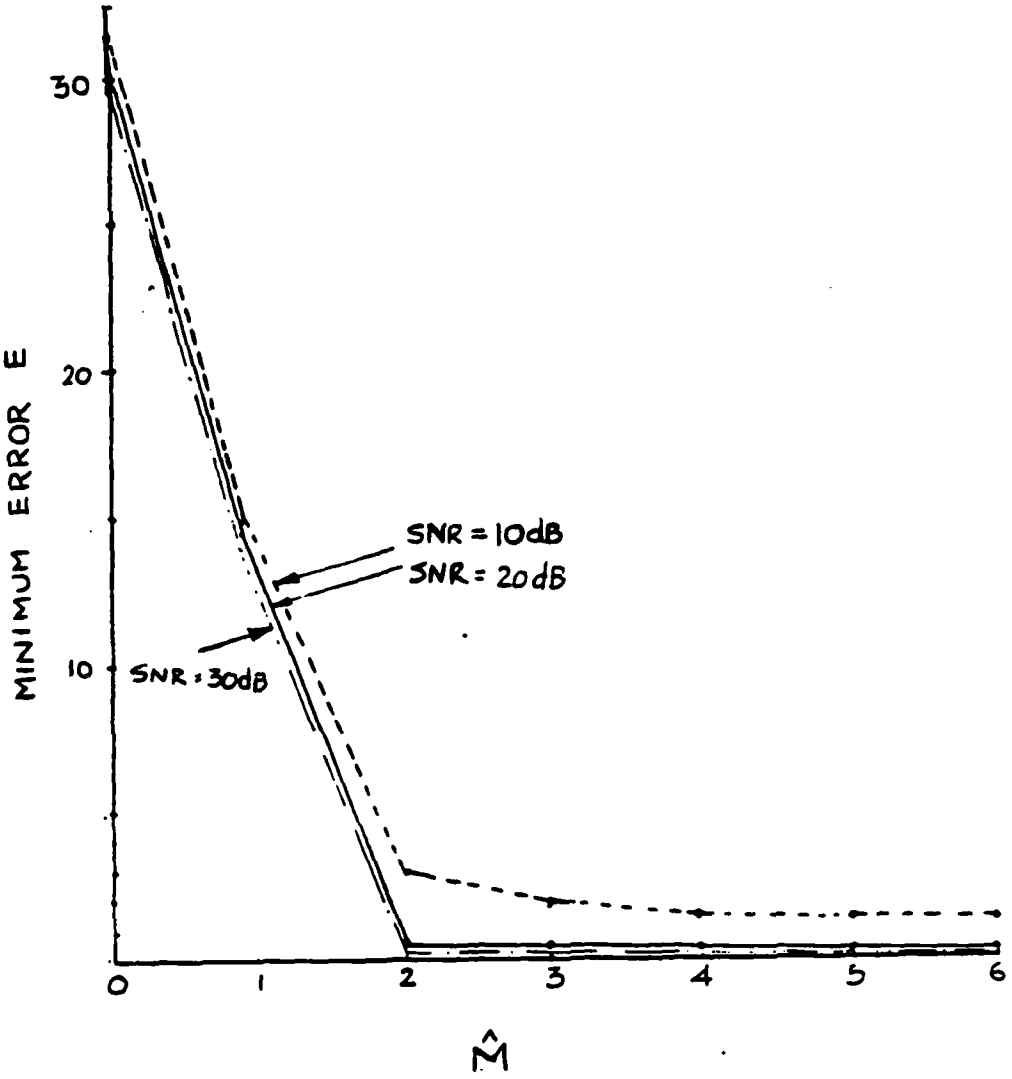


Figure 3.6: Minimum value of the error E in equation 3.7 for different values of M for a typical realization of the data is plotted at three SNR values.

CHAPTER 4: SINGULAR VALUE DECOMPOSITION AND IMPROVED PARAMETER ESTIMATION

In this chapter, we shall present two related methods, called the Tufts-Kumaresan (TK) method and the Improved Pisarenko (IP) method for estimating the signal parameters and M. These methods tend to be slightly more accurate than the modified Prony (MP) method presented in chapter 3.

The improvement in accuracy of the two methods presented in this chapter is attributable to the following two reasons. Firstly, as in the previous chapter, we trade redundancy in the degree, L of G(z), for increased tolerance to noise in the data. Secondly, we use an adaptive filtering method to improve the SNR before estimating the signal parameters. Motivation for the second point is as follows. Recall that in chapter 2, we showed that the parameters $\{s_k\}$ can be found exactly from the zeros of an L th degree polynomial G(z) if the N data samples are noiseless and L obeys certain inequalities. Therefore, one might expect that improving the SNR in the data and emulating the noiseless situation as closely as possible, one can improve the accuracy of the parameter estimates. But how does one improve the SNR in the data when the data record is short and nothing is known about the signal parameters $\{s_k\}$ and $\{a_k\}$? Although, we do not know the value of the signal parameters, a special property of exponential signals can be made use of for improving the SNR. For exponential signals, as we saw in chapter 2, the data matrix A'_f (or A'_b or A'_{fb} as the case may be) has a rank M when the data is noiseless. We might call the signals which have this property low rank or singular signals. This information can be used to get rid of some noise in the data. This is done by using SVD

of the data matrix. Once this is done, we can compute a polynomial G(z) as before, and obtain more accurate estimates of $\{s_k\}$.

The definition of SVD of a matrix and reasons for using it in our methods are given in section 4.1. In the TK method we show how to compute the coefficients of the polynomial G(z) using SVD. This is described in section 4.2.

Several researchers [64-77] in the past have used eigenvalue-eigenvector decomposition of autocorrelation matrices to estimate certain parameters. These works are closely related to our methods presented in this chapter. These relationships are spelled out in section 4.5. Of particular importance, from a historic point of view, is the work of Pisarenko [71]. He discovered that the frequencies of M undamped sinewaves can be found from the roots of an M th degree polynomial, formed with the elements of an eigenvector corresponding to the minimum eigenvalue of an (M+1) x (M+1) true autocorrelation matrix of data. Although, Pisarenko seems to have pointed out this fact first, the procedure he outlined for estimating the frequencies from a finite data record gives very poor estimates [45]. In section 4.3 we show how to improve upon his idea and how it relates to the TK method. In section 4.4 we discuss the trade-offs involved in choosing a value for the degree of G(z), for a given value of N, in order to obtain the most accurate parameter estimates.

4.1: REASONS FOR USING SVD OF A DATA MATRIX:

SVD is the generalization of eigenvalue decomposition to include rectangular matrices. Let A be a (mxn) matrix. The SVD of A is a product of three matrices [61,51]

$$A = U\Sigma V^{\dagger} \qquad 4.1$$

The matrices U and V are unitary [79] and Σ is a rectangular diagonal matrix of the same size as A with real, non-negative diagonal entries. The diagonal

entries σ_1, σ_2, σ_3,..... , called the singular values of A, are conventionally ordered with the largest in the upper left hand corner. The singular values of A are the non-negative square roots of the eigenvalues of $A^{\dagger}A$ or $A\,A^{\dagger}$.

In our case the A matrix is one of the data matrices, A_f, A_b, or A_{fb} rewritten here for convenience.

$$A_f = \begin{bmatrix} y(L) & y(L-1) & .... & y(1) \\ y(L+1) & y(L) & .... & y(2) \\ . & . & .... & . \\ . & . & .... & . \\ . & . & .... & . \\ y(N-1) & y(N-2) & .... & y(N-L) \end{bmatrix} \qquad 4.2$$

$$A_b = \begin{bmatrix} y^*(1) & y^*(2) & .... & y^*(L) \\ y^*(2) & y^*(3) & .... & y^*(L+1) \\ . & . & .... & . \\ . & . & .... & . \\ . & . & .... & . \\ y^*(N-L) & . & .... & y(N-1) \end{bmatrix} \qquad 4.3$$

$$A_{fb} = \begin{bmatrix} \frac{A_f}{A_b} \end{bmatrix} \qquad 4.4$$

The matrix A_f' (A_b' and A_{fb}') is the same as the above A_f (A_b and A_{fb}) except for an additional column to its left. (See section 2.1.) In our estimation algorithms one of the above matrices will be used depending on the type of signal in the data. If the signal is composed of exponentially damped sinusoids we may choose to use A_b (or A_b') and if the signals are undamped sinusoids, A_{fb} (or A_{fb}') will be used. We shall use the matrix symbol A to denote any of the above data matrices A_f, A_b, or A_{fb}, and A' to denote A_f', A_b', or A_{fb}'.

An important property of matrix A (or A'), is that it has a rank M if L, which determines the number of columns and rows of A (or A'), is chosen within the limits specified in chapter 2 (i.e., $M \le L \le N-M$ for A_f, A'_f, A_b, and A'_b and $M \le L \le (N-M/2)$ for A_{fb} and A'_{fb}) and the data is noiseless. Therefore $A^\dagger A$ (or $A'^\dagger A'$) and $A\,A^\dagger$ (or $A'\,A'^\dagger$) have M non-zero eigenvalues. Since the singular values of A (or A') are the (positive) square root of these eigenvalues, A (or A') has M non-zero singular values. We shall call the singular values of A $\sigma_1 \ge \sigma_2 \ge \ldots \ge \sigma_M$, $\sigma_{M+1}=0$, $\sigma_{M+2}=0\ldots$. The number of zero singular values depends on the order of A (or A').

Another important property of A (or A') regards the eigenvectors of $A^\dagger A$ (or $A'^\dagger A'$). The M principal eigenvectors of $A^\dagger A$ are linear combination of M linearly independent vectors $\underline{c}_i$, i=1,2,...M, where

$$\underline{c}_i = [1,\ e^{-s_i^*},\ e^{-2s_i^*},\ \ldots\ e^{-(L-1)s_i^*}]^T \qquad i=1,2,\ldots M \qquad 4.5$$

This fact is established in Appendix B. These vectors are said to span the signal subspace. The reason is that any L contiguous (complex conjugated) samples (in the reversed time direction) of the signal sequence can be represented as a linear combination of $\underline{c}_i$'s. The other eigenvectors of $A^\dagger A$ (or $A'^\dagger A'$) are orthogonal to the M principal eigenvectors (since $A^\dagger A$ (or $A'^\dagger A'$) is Hermitian) and hence orthogonal to the $\underline{c}_i$'s. This is another useful property of the eigenspace of such matrices.

If the data is noisy, the above properties are not exactly true, but they tend to be approximately true. That is, the M principal singular values of a noisy matrix A (or A') tend to be larger than the rest σ_{M+1}, σ_{M+2}, ... (which were originally zero). Also, the M eigenvectors which correspond to the M principal eigenvalues of $A^\dagger A$ or $A\,A^\dagger$ are less susceptible to noise perturbations in comparison to the rest of the eigenvectors. These statements are justifiable, qualitatively at least, using the first order perturbation results on eigenvalues

and eigenvectors of matrices due to Wilkinson [91]. In short, Wilkinson's results show that the amount of perturbation in an eigenvector direction (compared to the noiseless eigenvector direction) is dependent on the noise level and the separation of the corresponding eigenvalue from the rest. Therefore, the M principal eigenvectors of $A^{\dagger}A$ (or $A'^{\dagger}A'$) are much less perturbed since the corresponding eigenvalues are larger and generally well separated compared to the rest (whose eigenvalues were zero in the noiseless case). This robustness property of the principal eigenvectors is made use of in our improvements. Several authors have utilized this property of the principal eigenvectors in similar circumstances. See section 4.5 for details.

The above facts are related to a matrix approximation theorem due to Eckart and Young]83]. This connection was recognized by Holt and Anthill [76], Henderson [57], and Tufts et al. [56]. The Eckart and Young theorem states that the rank M matrix A_M minimizing the norm $||A-A_M||$ (defined as the trace of $(A-A_M)^{\dagger}(A-A_M)$) is given by

$$A_M = U\Sigma_M V^{\dagger} \qquad 4.6$$

where Σ_M is the matrix Σ with the singular values σ_{M+1}, σ_{M+2}, ... set to zero. Note that the matrix A_M is a sum of M outer products involving the M principal singular values and eigenvectors of $A^{\dagger}A$ and $A\,A^{\dagger}$. Thus, it is satisfying to note that the use of M principal singular values of A and the corresponding principal eigenvectors of $A^{\dagger}A$ (and $A\,A^{\dagger}$) directly follows from a (reduced rank) matrix approximation result. Recalling that the signal-only data matrix had a rank M, the rank M approximation (A_M) to the noisy data matrix A (or A') can be expected to be 'closer' to the signal-only data matrix than A. We shall now present a computer simulation which justifies this statement.

We construct a data matrix A with data samples consisting of two complex sinewaves and noise. Its rank M(=2) approximation, A_M is found by using SVD

as explained above. Another data matrix $\bar{A}$ is constructed with signal only. Its rank, of course, is 2. Then we measure the closeness of the matrices A and A_M to $\bar{A}$ by computing the following quantity Q.

$$Q = \|A^\dagger A - \bar{A}^\dagger \bar{A}\| - \|A_M{}^\dagger A_M - \bar{A}^\dagger \bar{A}\|$$

Thus if Q is positive it means that the matrix A_M is closer to $\bar{A}$ than A. Five hundred such trials are performed with d'fferent noise epochs and the histogram of Q is plotted in figures 4.1 and 4.2 at two different SNR values. Even at 0 db SNR, Q is almost always positive, showing that the matrix A_M is almost always closer to the noiseless data matrix $\bar{A}$. In other words, computing A_M (or using the M principal eigenvalues and eigenvectors of $A^\dagger A$) indeed enhances the signal. See also ref. [56,50].

Based on the above justifications the following statements can be made.

(1) From the size of the singular values of A, an estimate of M, namely $\hat{M}$ can be obtained. This is because the principal singular values tend to be larger than the noise related singular values which were originally zero.

(2) The M principal eigenvectors approximately span the signal subspace and hence using them as opposed to the whole A (or A') matrix amounts to an effective SNR improvement [56].

It may be noted that SVD is used here not to improve the numerical stability of computations for which it is often intended [51,61], but to improve the SNR in the data. Equivalently, the eigenvalue decomposition of $A^\dagger A$ or $A\,A^\dagger$ can be used with little loss in accuracy.

4.2: TUFTS-KUMARESAN METHOD:

In this method we show how to take advantage of the properties of the principal eigenvalues and eigenvectors to compute a polynomial G(z) which results in improved estimation accuracy. We set up as in the modified Prony (MP) method (in chapter 3) the following set of equations using the noisy data samples

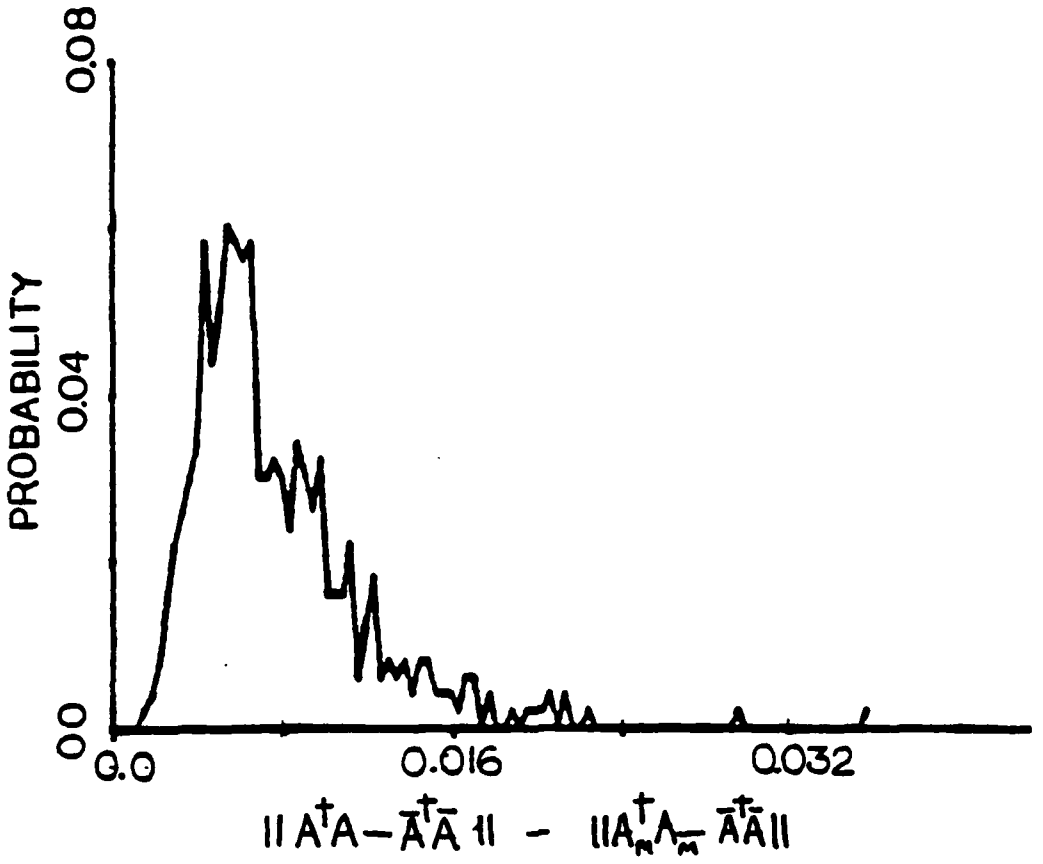


Figure 4.1: Histogram of $Q = \|A^\dagger A - \bar{A}^\dagger A\| = \|A_M^\dagger A_M - \bar{A}^\dagger \bar{A}\|$. A is the data matrix with noisy data samples. The same formula used for data generation in figure 3.1 is used. 500 independent trials were performed. $\bar{A}$ is a data matrix with noiseless data samples. A_M is the rank 2 approximation to A. $\|X\|$ is the trace of $X^\dagger X$. L is chosen 12. N=25. SNR=20 db.

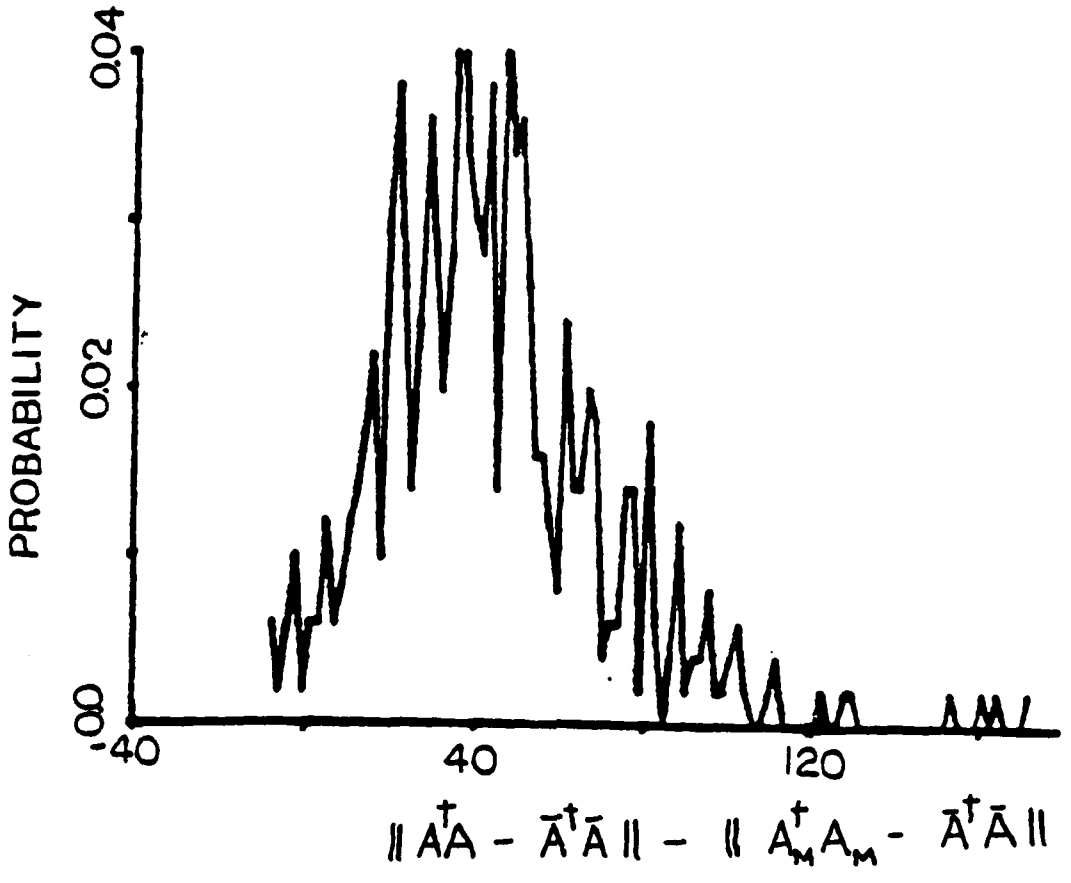


Figure 4.2: Histogram of Q, SNR=0 db.

y(n),n=1,2,...N.

$$A\,\underline{g} \cong -\underline{h} \tag{4.7}$$

Unlike the MP method we do not attempt to satisfy the above equations in the least square sense directly. First, we improve the SNR in the data as described above by computing the SVD of A. Based on the size of the singular values of A we decide on the value or $\hat{M}$. A $\hat{M}$ which approximates A in the least square sense and has a rank $\hat{M}$ is given by

$$A_{\hat{M}} = U\Sigma_{\hat{M}}V^{\dagger} \tag{4.8}$$

Then, we solve the following linear equations instead of the original equation in 4.7.

$$A_{\hat{M}}\underline{g} = -\underline{h} \tag{4.9}$$

There are many possible solutions to these equations because the rank of $A_{\hat{M}}$ is less than full. But using the pseudoinverse of $A_{\hat{M}}$, we find the vector g as follows.

$$\underline{g} = -A_{\hat{M}}^{\#}\underline{h} \tag{4.10}$$

This pseudoinverse solution has a desirable property. That is, it finds a solution vector $\underline{g}$, which has the minimum length $\|\underline{g}\|$. Note that this is the condition we imposed on the $\underline{g}$ vector in chapter 2 to restrain the L-M extraneous zeros to be inside the unit circle. The pseudoinverse of $A_{\hat{M}}$ can be written as [61],

$$A_{M}^{\#} = U\Sigma_{M}^{\#}V^{\dagger} \tag{4.11}$$

where $\Sigma_{\hat{M}}^{\#}$ is obtained from $\Sigma_{\hat{M}}$ by replacing each positive diagonal element by its reciprocal. Then the minimum norm $\underline{g}$ vector can be written as,

$$\underline{g} = -\sum_{k=1}^{\hat{M}} \frac{1}{\sigma_k}(\underline{v}_k^{\dagger}\underline{h})\underline{u}_k \tag{4.12}$$

where $\underline{v}_k$ and $\underline{u}_k$, k=1,2,..... are the columns of V and U. The minimum norm vector $\underline{g}$ can also be written as

$$\underline{g} = -\sum_{k=1}^{\hat{M}} \frac{1}{\lambda_k} (\underline{u}_k^{\dagger} A^{\dagger} \underline{h}) \underline{u}_k \qquad 4.13$$

where λ_k's are the eigenvalues of $A^{\dagger}A$ or $A\,A^{\dagger}$.

Thus in the TK method, we try to emulate the noiseless case by first finding a matrix $A_{\hat{M}}$ which is as close as possible to the noiseless signal matrix and then finding a coefficient vector $\underline{g}$ which has minimum norm. If $\hat{M}$ is chosen to be equal to or slightly greater than M and the data SNR is high enough, the behaviour of the signal and extraneous zeros of G(z) obtained by the TK method are close to that predicted in the noiseless case in chapter 2.

Figure 4.3 shows the zeros of the polynomial G(z) obtained by using the TK method for several trials. The coefficients of G(z) are calculated as in equation 4.12. The same data sets and value of L used in figure 3.3 are used. For comparison, the zeros of G(z) when the data is noiseless is plotted in figure 4.4. Comparing figures 3.3, 4.3, and 4.4, it should be clear that the effect of using SVD is to enhance the SNR and emulate the noiseless case as closely as possible. Figure 4.5 (compare this with figure 3.5) shows that situation in which the signal is composed of exponentially damped signals. The equations $A_b\underline{g}=-\underline{h}_b$ were used in this case.

Numerical analysts, attempting to solve ill-conditioned linear equations such as here, have routinely used orthogonalizaton techniques such as QR decomposition [61] or truncated SVD as above, to solve them. But our situation is slightly different. In our case, in addition to using SVD of A, we can choose the value of L (in other words we can alter the set of equations $A\underline{g}=-\underline{h}$ itself), which, as we saw in the previous chapter, has considerable effect on the improvement in estimation accuracy. The choice of the value of L is discussed in section 4.4.

We may point out that Strand [80] has proposed a method to compute the truncated SVD solution as in 4.12 without having to compute the SVD explicitly.

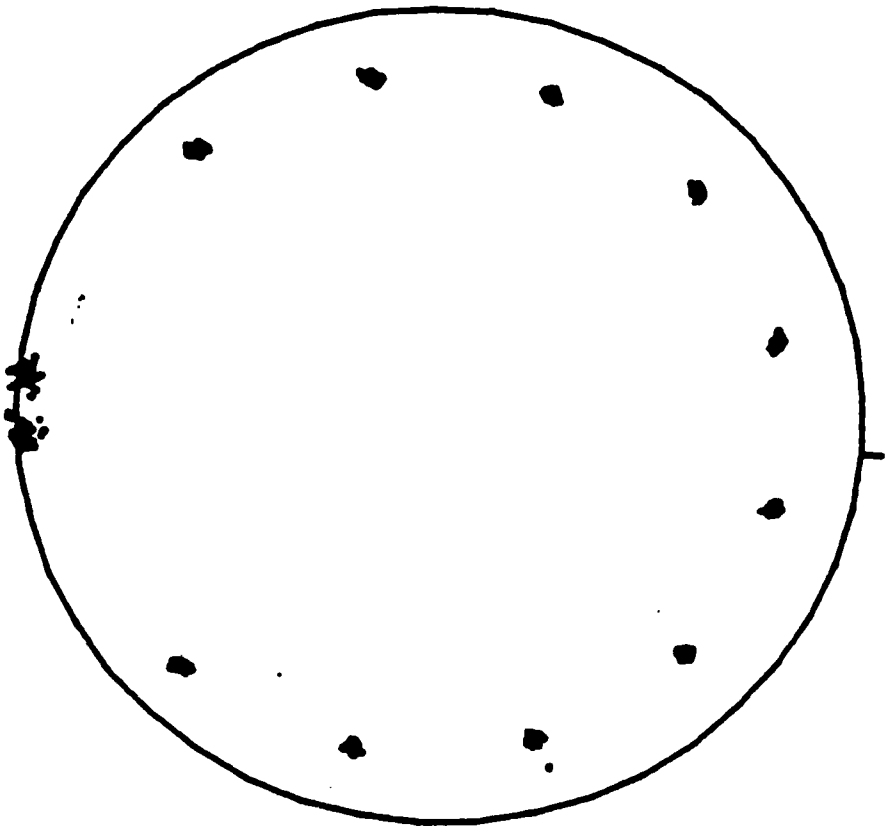


Figure 4.3: The zeros of G(z) for 50 independent trials. $\underline{g}$ vector is calculated as in equation 4.12 with M=2. The forward-backward prediction equations were used. The same data sets used in figure 3.3 are used.

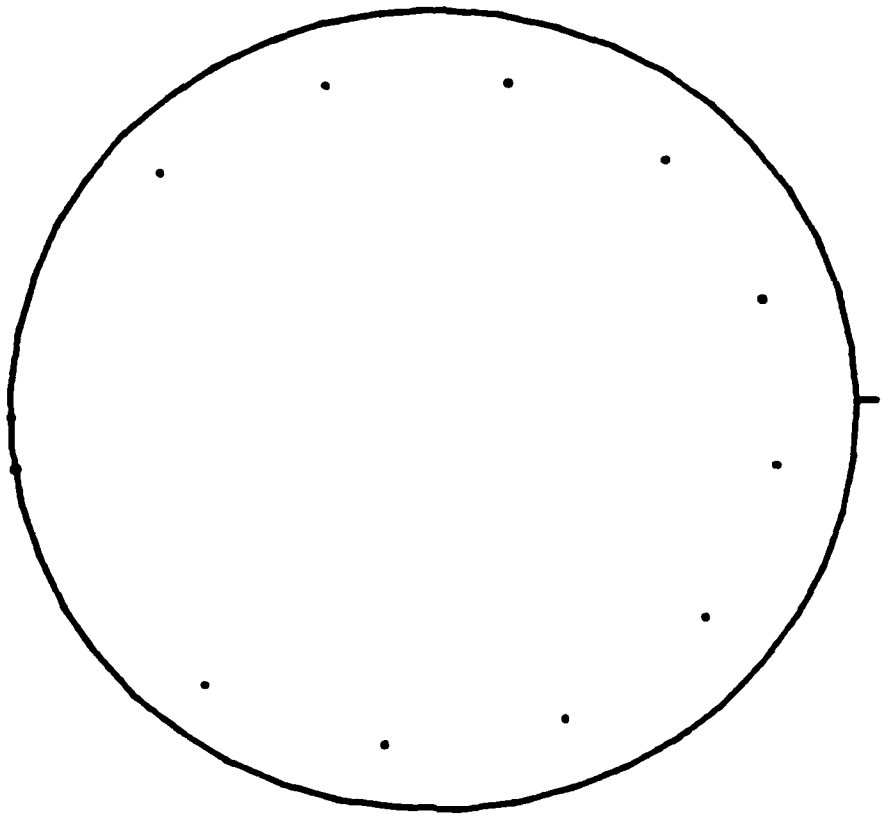


Figure 4.4: The zeros of G(z) for noiseless data. $\underline{g}$ vector is calculated as in equation 4.12. M=2.

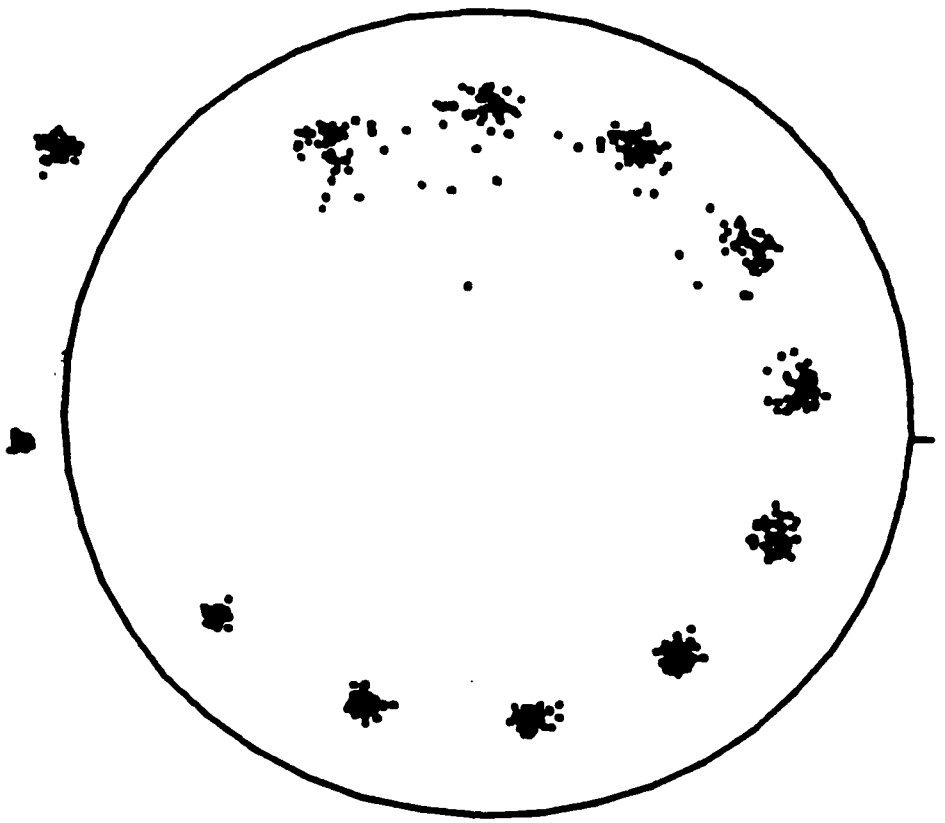


Figure 4.5: The zeros of G(z) for exponentially damped sinusoidal signals in noise. The same data sets used in fig. 3.5 are used. The equation $A_b \underline{g} \approx -h_b$, is used.

4.3: IMPROVED PISARENKO METHOD:

In this section also, we use SVD of a data matrix (A'), but the emphasis is on using the smaller singular values and corresponding eigenvectors. We improve upon an idea proposed by Pisarenko [71] and hence the name improved Pisarenko method. It is not significantly different from the TK method philosophically. But, it helps us point out the links between several eigenvector methods that are available today (August 1982). These links are described in section 4.4.

4.3.1: PISARENKO METHOD AND ITS DIFFICULTIES:

Pisarenko [71] seems to have originated the idea of using the minimum eigenvalue-eigenvector polynomial to find the frequencies and power of M undamped sinewaves in white noise. By 'eigenvector polynomial' we mean the following. If $\underline{e}=(e_0,e_1.....e_L)^T$ is an (eigen)vector, then the corresponding eigenvector polynomial is defined as $\sum_{k=0}^{L} e_k z^{-k}$. Pisarenko, using the elements of the eigenvector corresponding to the minimum eigenvalue of an (M+1)x(M+1) true auto correlation matrix of data consisting of M complex sinewaves and white noise, constructed an M th degree polynomial and showed that the roots of this polynomial (which lie on the unit circle) gave the frequencies of the sinewave.

There are at least three difficulties associated with the Pisarenko method. (1) The actual algorithm he devised for finite data records does not utilize the data effectively to obtain accurate estimates. One reason is the limitation on the order of the matrix, which has to be (M+1)x(M+1). This causes eigenvector polynomial G(z) to be of degree M. We know from the previous chapter and intuition that the degree L of G(z) needs to be greater than M to improve the accuracy of the parameter estimates. (2) His algorithm can only be used for estimating the frequencies of undamped sinewaves. This is because he forces the estimated autocorrelation matrix to be Toeplitz. The

effect of this is, that the polynomial constructed with the elements of the eigenvector corresponding to the minimum eigenvalue (if it is unique) of the Toeplitz matrix will have zeros on the unit circle [81]. Obviously, zeros of such a polynomial cannot be used to estimate the parameters $\{s_k\}$ of damped sinusoidal signals. (3) Also, Pisarenko's method will give, in general, biased values of the frequencies of even noiseless, undamped sinewaves for finite data records. The last two problems are caused by forcing the estimated autocorrelation matrix to be Toeplitz.

4.3.2: BASIS FOR IMPROVING PISARENKO'S METHOD:

The following observations form the basis of our improvements to Pisarenko's method. Recall that in chapter 2 (see propositions 1, 2, and 3) we pointed out that, if the data is noiseless and if a vector $\underline{g}'$ (g_0, the first element need not be unity) satisfies the homogeneous equation,

$$A'\underline{g}'=\underline{o} \qquad 4.14$$

then the signal parameters can be exactly recovered from the zeros of $G(z) = \sum_{k=0}^{L} g_k z^{-k}$. Premultiplying the above equation by $A'^{\dagger}$, we get

$$A'^{\dagger}A'\underline{g}'=\underline{0}$$

Since the rank of A' is M (see chapter 2), $A'^{\dagger}A'$ will have only M non zero eigenvalues. Therefore, the above is an eigen equation corresponding to a zero eigenvalue (which has a multiplicity L+1-M) and $\underline{g}'$ is an eigenvector. It is possible to find L+1-M such eigenvectors. We shall say that these vectors span the null or noise subspace. $\underline{g}'$ can be chosen as any vector lying in this subspace. If L is chosen equal to M, then $\underline{g}'$ will be unique. The maximum value of L can be N-M or N-M/2 depending on the type of matrix used (see chapter 2, section 2.1).

This is to be contrasted with the Pisarenko method where the Toeplitz correlation matrix R, with elements $r_i = \frac{1}{N}\sum_{k=1}^{N-|i|} (y(k)\, y^*(k-i))$, $|i| \leq M+1$, is

used. The eigenvector polynomial corresponding to the minimum eigenvalue of R, even if the data is noiseless, will not give the exact value of $\{s_k\}$ for finite data records, in general. Therefore, we shall use in our improvements the correlation matrix $A'^{\dagger}A'$ instead of R since its eigenvector polynomial gives the exact parameters, at least when the data is noiseless, independent of the type of the signal.

Secondly, as we saw earlier, a polynomial G(z) with degree L>M tends to give more accurate parameter estimates. This turns out to be true in this case also. However, the value of L determines the order of the (L+1)x(L+1) matrix $A'^{\dagger}A'$ and hence the number of eigenvectors (L+1-M) in the null subspace or noise subspace of $A'^{\dagger}A'$. The questions are which of these vectors should be used to construct a polynomial G(z)? Can they be combined linearly (if yes, how?) to form the coefficient vector $\underline{g}'$?

4.3.3: HOW TO CONSTRUCT A POLYNOMIAL G(z) IN IP METHOD:

Let us write the SVD of A' as follows.

$$\underset{m \times (L+1)}{A'} = \underset{(m \times m)}{P} \; \underset{m \times (L+1)}{\Lambda} \; \underset{(L+1 \times L+1)}{Q^{\dagger}} \qquad 4.16$$

where Λ is the rectangular diagonal matrix with singular values along the diagonal which are the positive square roots of the eigenvalues of $A'^{\dagger}A'$ or $A'\,A'^{\dagger}$. The value of m will be determined by the type of data matrix used. That is, for A'_{fb}, m=2(N-L) and for A'_f and A'_b, m=(N-L). P and Q are unitary matrices. The columns of P and Q, $\underline{p}_i$, i=1,2,...m and $\underline{q}_i$, i=1,2,...L+1 are the eigenvectors of $A'\,A'^{\dagger}$ and $A'^{\dagger}A'$, respectively.

Any one of the eigenvectors in the noise subspace $\underline{q}_{M+1}, \underline{q}_{M+2}, \ldots \underline{q}_{L+1}$ can be used to construct an eigenvector polynomial. Examples are given in figure 4.6-4.9 where different noise subspace eigenvector polynomial zeros are plotted for 50 different trials. Note that here too (as in the modified Prony method), we

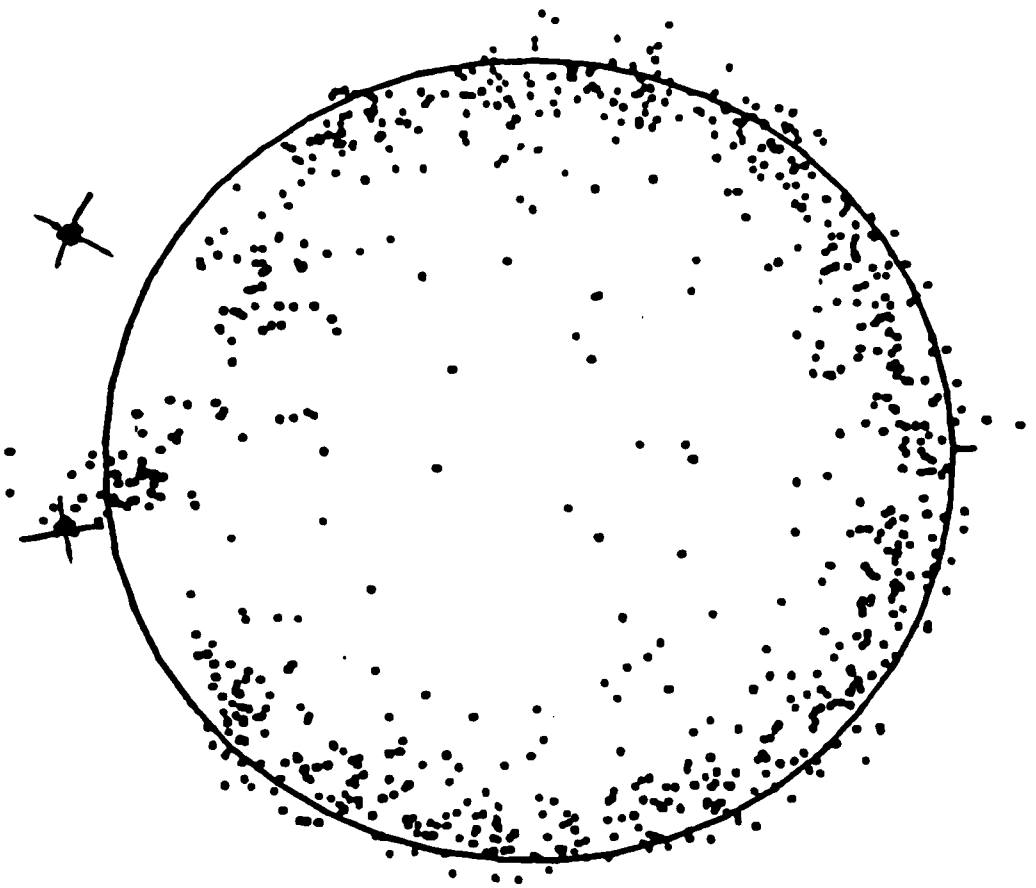


Figure 4.6: The zeros of a noise subspace eigenvector polynomial $Q(z) = q_{L+1,k} z^{-k}$. $\underline{q}_{L+1} = (q_{L+1,1}, q_{L+1,2}, \cdots\cdots q_{L+1,L+1})^T$. $\underline{q}_{L+1}$ is the L+1 st eigenvector of $A'^{\dagger}_b A'_b$. SNR=30 db. L=18, the same data sets in figure 3.4 are used in figures 4.6-4.10.

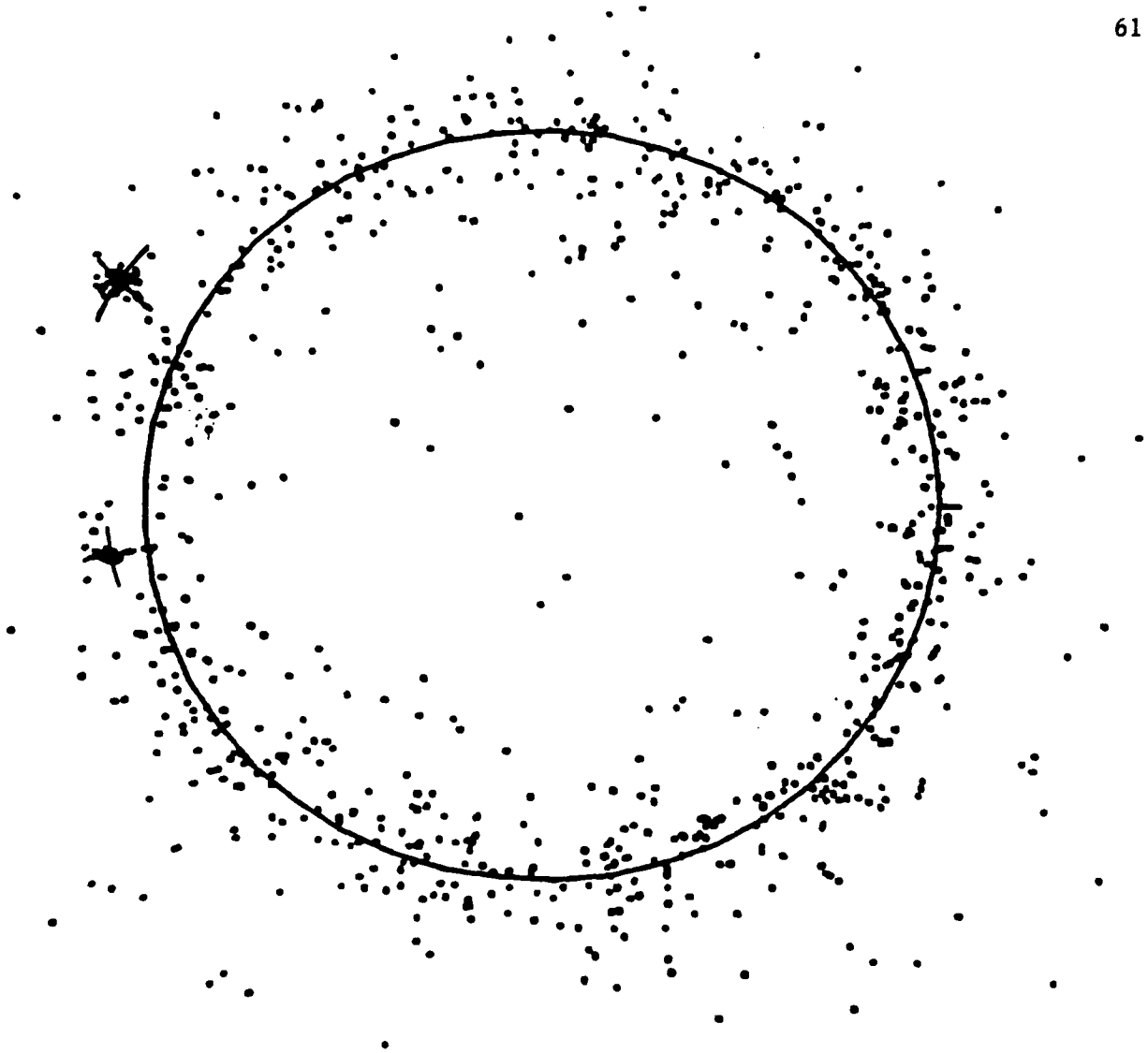


Figure 4.7: The zeros of an eigenvector polynomial. The eigenvector $\underline{q}_6$ is used.

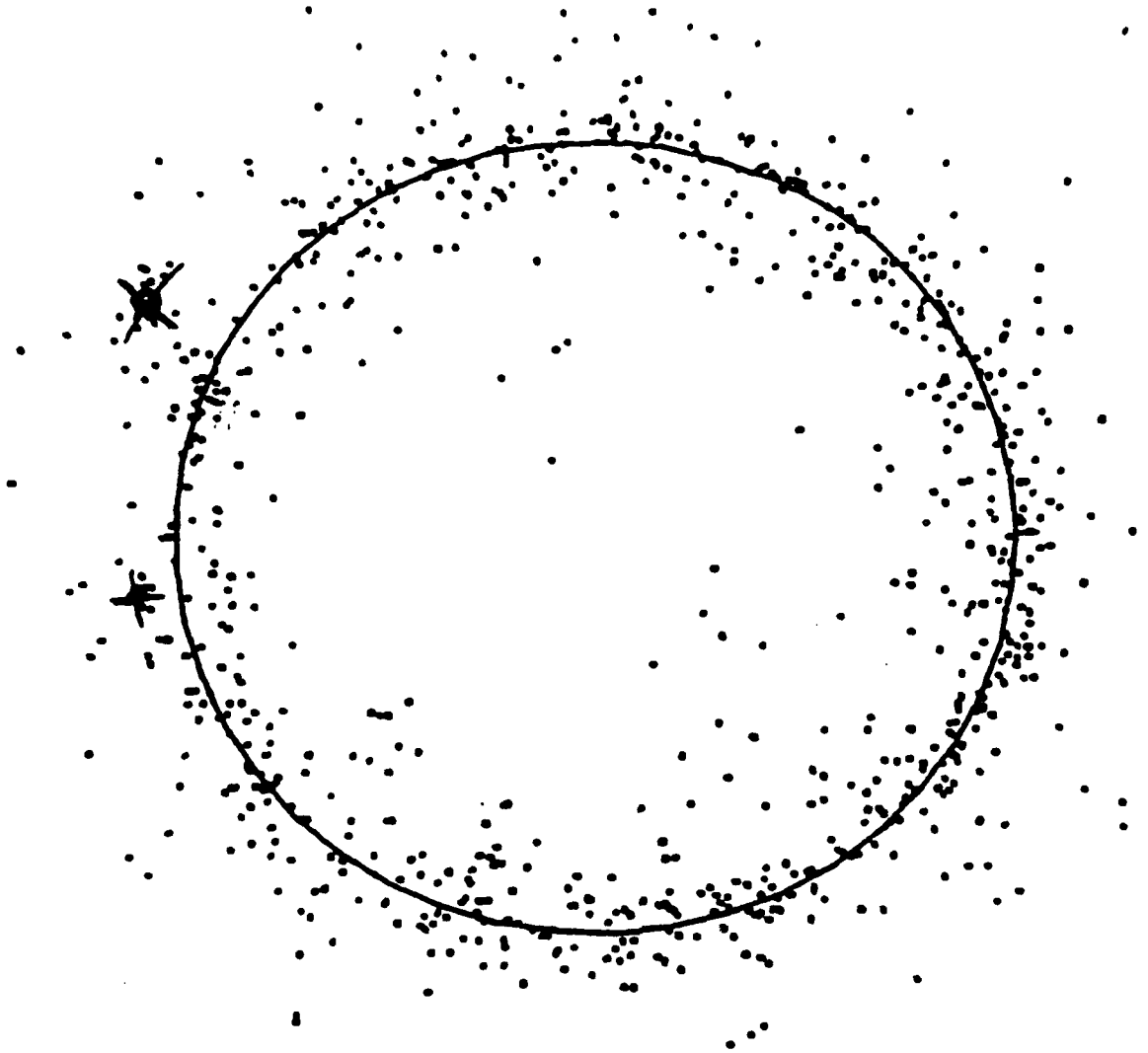


Figure 4.8: The zeros of an eigenvector polynomial. The eigenvector q_{12} is used.

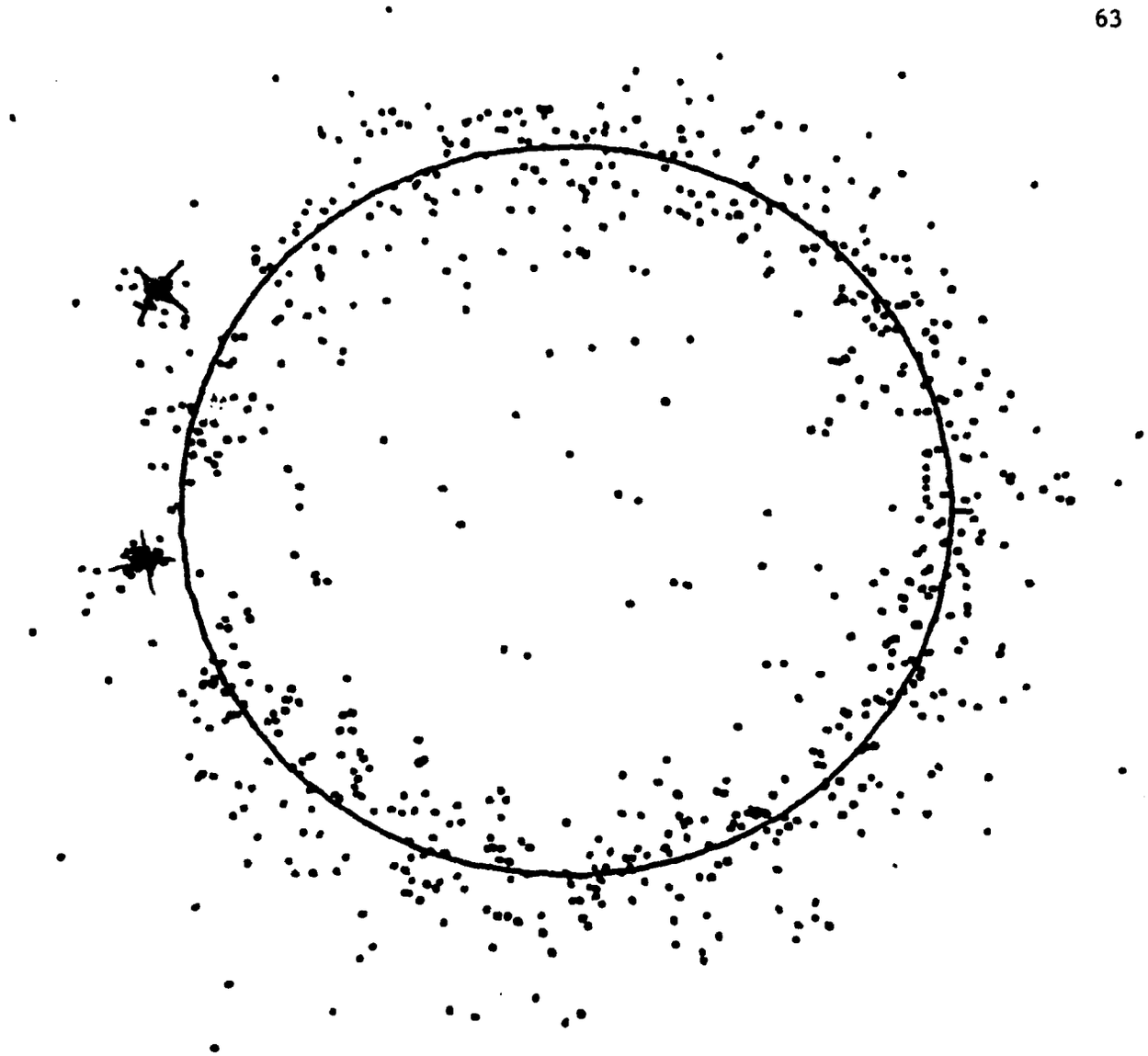


Figure 4.9: The zeros of an eigenvector polynomial. The eigenvector $\underline{q}_{16}$ is used.

face the same problem of distinguishing the signal zeros from the extraneous ones. The procedure we outlined in section 3.2.2 can be used to distinguish the L-M extraneous zeros of these eigenvector polynomials.

Instead, since each noise subspace eigenvector is a potential candidate from which the signal parameters can be estimated, they can be combined to form a single vector $\underline{g}'$. That is, $\underline{g}'$ can be formed as a linear combination of $\underline{q}_{M+1}, \underline{q}_{M+2}, \cdots \underline{q}_{L+1}$

$$\underline{g}' = \sum_{k=1}^{L+1-M} b_k \, \underline{q}_{k+M} \qquad 4.17$$

How do we choose the b_k's? There are several choices. For example, $b_k=1$, $k=1,2,\ldots$ L+1-M or the b_k's can be set equal to the singular values $\delta_{M+1}, \delta_{M+2}, \ldots \delta_{L+1}$ etc.

Another alternative is to find the b_k's by imposing the constraints used in chapter 2; i.e., $g_0=1$ and $|g_1|^2 + |g_2|^2 + \ldots + |g_L|^2$ is minimum. Then, finding the b_k's subject to these constraints is a standard minimization problem [82]. But the same vector $\underline{g}'$ can be found more easily using the following observation. $\underline{g}'$ must be orthogonal to the M principal eigenvectors $\underline{q}_1, \underline{q}_2, \ldots \underline{q}_M$ of $A'^{\dagger}A'$ because it lies in the span of $\underline{q}_{M+1}, \underline{q}_{M+2}, \ldots \underline{q}_{L+1}$. Therefore, the following inner products can be set to zero.

$$\underline{q}_i^{\dagger} \, \underline{g}' = 0 \qquad i=1,2,\ldots M \qquad 4.18$$

Let us call the j th element of the i th eigenvector q_{ij}. We can rewrite the above equations as follows.

$$\begin{bmatrix} q_{11}^* & q_{12}^* & \cdots\cdots & q_{1\,L+1}^* \\ q_{21}^* & q_{22}^* & \cdots\cdots & q_{2\,L+1}^* \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ q_{M1}^* & q_{M2}^* & & q_{M\,L+1}^* \end{bmatrix} \begin{bmatrix} 1 \\ g_1 \\ g_2 \\ \\ g_L \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix} \qquad 4.19$$

Taking the first column of the above matrix to the right side we can rewrite the equation as,

$$\begin{bmatrix} q_{12}^* & q_{13}^* & \cdots\cdots & q_{1L+1}^* \\ q_{22}^* & q_{23}^* & \cdots\cdots & q_{2L+1}^* \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ q_{M2}^* & q_{M3}^* & & q_{ML+1}^* \end{bmatrix} \begin{bmatrix} g_1 \\ g_2 \\ \cdot \\ \cdot \\ g_L \end{bmatrix} = \begin{bmatrix} q_{11}^* \\ q_{21}^* \\ \cdot \\ \cdot \\ q_{M1}^* \end{bmatrix} \qquad 4.20$$

and in matrix form as, $(\underline{g}'=(1,\underline{g}^T)^T)$,

$$C\,\underline{g} = -\underline{c} \qquad 4.21$$

It might be useful to weight the above equations 4.19 based on the size of the corresponding singular values. That is, the first equation is multiplied by the first singular value and so on. The reason is that the eigenvectors corresponding to the larger singular values tend to be less noisy. Since L>M, the above system has more unknowns than equations. The minimum norm solution (which minimizes the quantity $|g_1|^2$, + $|g_2|^2$ .... + $|g_L|^2$ as desired) can be found using the pseudo-inverse of C.

$$\underline{g} = -C^\dagger[CC^\dagger]^{-1}\underline{c} \qquad 4.22$$

But $CC^\dagger$ is $I-\underline{c}\,\underline{c}^\dagger$ which follows from the orthonormality of the eigenvectors of $A'^\dagger A'$. Therefore the expression for $\underline{g}$ can be simplified to

$$\underline{g} = \frac{C^\dagger \underline{c}}{1-\|\underline{c}\|^2} \qquad 4.23$$

Thus the desired $\underline{g}$ vector can be obtained from the M principal eigenvectors of $A'^\dagger A'$ by linearly combining the last L elements of each vector. An equivalent expression for $\underline{g}$ can be obtained by using the L+1-M noise subspace eigenvectors as well. Figure 4.10 shows the zeros of a polynomial formed with the $\underline{g}$ vector computed as per equation 4.23. Note that the effect of using the said constraints is to make the L-M extraneous zeros fall inside the unit circle.

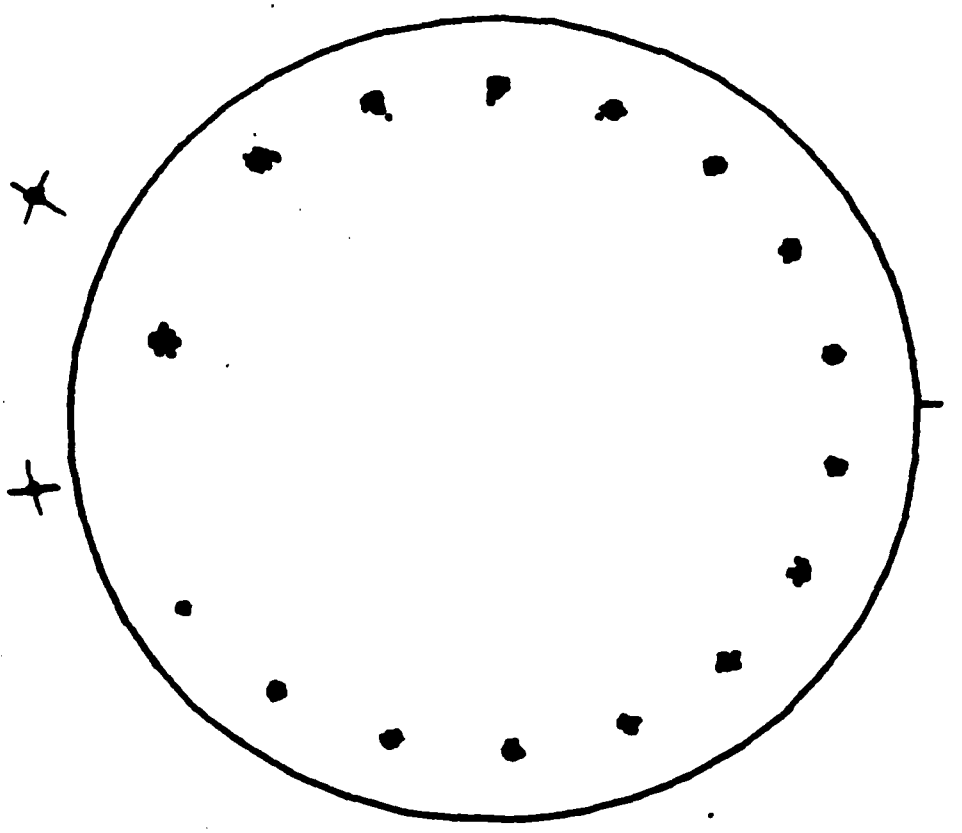


Figure 4.10: The zeros of G(z) obtained by IP method. In this case, the $\underline{g}$ vector was calculated as in equation 4.23. That is $\underline{g}'$ lies in the space of all the L+1-M noise subspace eigenvectors of $A_b'^{\dagger}A_b'$. Also the coefficients of G(z) are constrained to be $g_0=1$, $|g_1|^2 + |g_2|^2 + \ldots\ldots + |g_L|^2$ is a minimum.

4.3.4: SUMMARY:

Any eigenvector corresponding to the L+1-M smaller eigenvalues of $A'^{\dagger}A'$ is orthogonal to the signal subspace approximately. Therefore, a polynomial constructed with any one of the L+1-M eigenvectors will have zeros (close to) e^{s_k} (or $e^{-s_k^*}$, in the case when A_b' is used), k=1,2,...M. The L+1-M eigenvectors can be linearly combined to form a single vector $\underline{g}'$ in several ways. One particularly interesting way is by constraining the coefficients of G(z), which are the elements of $\underline{g}'$, to be $g_0=1$, and $|g_1|^2 + |g_2|^2 + \ldots + |g_L|^2$ to be minimum. This helps us identify the signal zeros from the extraneous zeros. Using these constraints makes the IP method resemble (in performance as well) the TK method closely. Some other choice for the b_k's which retains the 'homogenity' (see Henderson [57] for explanation) may be preferable. In this case, the elimination of the L-M extraneous zeros is a problem. It can be solved using the method in 3.2.2.

In the next section we discuss the choice of the value of L for a given number of data samples.

4.4: CHOICE OF THE VALUE OF L AND A SPECIAL CASE AT THE EXTREME VALUE OF L:

Choosing the best value of L, the degree of G(z), is crucial to achieving the best parameter estimation performance. The value of L, for a given value of N, determines the order of the data matrices A or A'. Earlier, in chapter 3, we observed that the parameter estimation accuracy increased with increase in the value of L. We also observed that L should be less than or equal to N-M (or N-M/2 in the case of forward-backward equations), where M is the number of signal components in the data. The interesting question, however, is how to choose the value of L for a given N to achieve the best possible performance. We were not able to obtain an expression for best L, in terms of N, M, SNR, and the signal parameters. However, to see the tradeoffs involved in the choice of L, we shall present a simple example.

4.4.1: AN EXAMPLE:

In this example we shall derive a simple expression for the first order perturbations on the signal zero of G(z) as a function of the perturbations on the coefficients of G(z), and the value of L. Recall that the signal parameter estimates are obtained from the locations of the signal zeros. First we shall calculate the coefficients of G(z) when the data is noiseless.

Let us consider the case when the data is a complex sinusoid.

$$y(n) = e^{jw_1 n} \qquad n=1,2,... \tag{4.24}$$

w_1 is the angular frequency of the sinusoid. We can write the linear prediction equations (for L>M=1) as,

$$\sum_{k=1}^{L} y(n-k)\, g_k = -y(n), \qquad n=L+1 \tag{4.25}$$

or in matrix form

$$A_f \underline{g} = -\underline{h}_f \tag{4.26}$$

Since the signal sequence has one complex sinusoid, one equation is necessary

and sufficient to find the coefficients g_k,k=1,2,...L. The only signal subspace eigenvector of $A^{\dagger}A$ can be shown to be

$$\underline{v}_1 = (1/L)^{\frac{1}{2}} (1, e^{jw_1}, e^{2jw_1}, \ldots e^{(L-1)jw_1})^T \qquad 4.27$$

Further the minimum norm solution vector $\underline{g}$ can be shown to be

$$\underline{g} = (-1/L) (e^{jw_1}, e^{2jw_1}, \ldots e^{Ljw_1})^T \qquad 4.28$$

Therefore the polynomial G(z) is

$$G(z) = 1 - \sum_{k=1}^{L} (1/L) e^{jw_1 k} a^{-k} \qquad 4.29$$

Let $z_1, z_2, \ldots z_L$ be the zeros of G(z). $z_1 = e^{jw_1}$ is the signal zero.

Now we can study the first order perturbation of the zeros of G(z) for incremental changes in its coefficients by using the formula (see for e.g. [20])

$$dz_i = \frac{-\sum_{k=1}^{L} z_i^{(1-k)} \Delta g_k}{\prod_{\substack{k=1 \\ k\neq i}}^{L} (1-z_k z_i^{-1})} \qquad 4.30$$

where Δg_k are the perturbations on the coefficients of G(z). If G(z) has tightly clustered zeros, small changes in g_k could cause large changes in the zeros of G(z) since the denominator will tend to be small. Note that the minimum norm coefficient vector we have used in our methods is advantageous in this respect also since it tends to distribute the zeros around the unit circle (see chapter 2) making the denominator large in the above formula. Specializing the above formula to the signal zero $z_1 = e^{jw_1}$, we get

$$dz_1 = \frac{-\sum_{k=1}^{L} z_1^{(1-k)} \Delta g_k}{G'(z) \text{ at } z=z_1} \qquad 4.31$$

where

$$G'(z) = \prod_{k=2}^{L} (1-z_k z_1^{-1})$$

$$= \prod_{k=0}^{L-1} \frac{L-k}{L} e^{jw_1 k} z^{-k} \qquad 4.32$$

Therefore, after some simplification dz_1 is given by the formula,

$$dz_1 = \frac{-2\ \underline{v}_1^{\dagger}\Delta\underline{g}}{L+1} \qquad 4.33$$

where

$$\Delta\underline{g} = (\Delta g_1,\ \Delta g_2,\ \ldots.\ \Delta g_L)^T \qquad 4.34$$

$\Delta\underline{g}$ is the vector of perturbations on the coefficient vector $\underline{g}$. Tha magnitude of dz_1 determines the perturbation of the zero z_1 and hence the accuracy of the estimate $s_1=jw_1$. There are two factors which determine $|dz_1|$.

(1) Clearly, as the value of L is increased $|dz_1|$ tends to be smaller, since L occurs in the denominator.

(2) But, $|\underline{v}_1^{\dagger}\Delta\underline{g}|$ which appears in the numerator tends to be larger as L is increased. To see the effect of L qualitatively, let us take the case of TK method. A similar case can be made for the IP method as well. Recall that in the TK method, the $\hat{\underline{g}}$ vector ($\hat{\underline{g}}$ is the same as $\underline{g}$ vector. The $\hat{}$ on matrices, vectors, and scalars is used to denote the presence of noise in the data (in this subsection only)) is calculated as (see equation 4.13)

$$\hat{\underline{g}} = -(1/\hat{\lambda}_1)\ (\hat{\underline{v}}_1^{\dagger}\hat{\underline{r}})\hat{\underline{v}}_1 = c_1\hat{\underline{v}}_1 \qquad 4.35$$

where $c_1 = -(\hat{\underline{v}}_1^{\dagger}\hat{\underline{r}})/\hat{\lambda}_1$, $\hat{\lambda}_k$, $\hat{\underline{v}}_k$ are the eigenvalues and eigenvectors of $\hat{A}^{\dagger}\hat{A}$ and $\hat{\underline{r}}=\hat{A}^{\dagger}\hat{\underline{h}}$. In the absence of noise in the data the $\underline{g}$ vector is given by

$$\underline{g} = -(1/\lambda_1)\ (\underline{v}_1^{\dagger}\underline{r})\underline{v}_1 = c_1\underline{v}_1 \qquad 4.36$$

where $c_1 = -(\underline{v}_1^{\dagger}\underline{r})/\lambda_1$.

Therefore

$$\Delta\underline{g}^{TK} = \hat{\underline{g}} - \underline{g} = \hat{c}_1\hat{\underline{v}}_1 - c_1\underline{v}_1 \qquad 4.37$$

Superscript 'TK' denotes $\Delta\underline{g}$ vector corresponding to TK method.

Premultiplying by $\underline{v}_1^{\dagger}$ we get

$$\underline{v}_1^{\dagger}\Delta\underline{g}^{TK} = \hat{c}_1(\underline{v}_1^{\dagger}\hat{\underline{v}}_1) - c_1, \qquad \|\underline{v}_1\| = 1$$

Therefore $\Delta\underline{g}^{TK}$ is primarily a function of noise in the first eigenvalue and

eigenvector of $A^{\dagger}A$. This is dependent on L because, for a fixed value of N, the perturbation in each element of $A^{\dagger}A$ is dependent on the value of L. For example, the (i,j) th element in $A_f A_f$ is $\sum_k y^*(k-i+1)y(k-j+1)$. Therefore the larger the value of L, smaller is the averaging and hence larger is the relative perturbation on λ_1 and $\underline{v}_1$, leading to larger $|dz_1|$.

Hence there are two factors working in the opposite direction as L is increased. Therefore, we can expect an optimum intermediate value of L in between M and N-M (or N-M/2).

4.4.2: A SPECIAL CASE: [92]

In the previous section, we argued that a compromise value of L in between M and N-M (or N-M/2) should be used to minimize the perturbation on the signal zero(s). However, a special case occurs at the extreme value of I (=N-M for matrices A_f and A_b and N-M/2 for A_{fb}). The advantage of choosing this particular value of L is that no SVD calculations are required in computing the coefficient vector $\underline{g}$. Also, for this extreme value of L the coefficient vectors $\underline{g}$ (minimum norm vector) given by the three methods (MP,TK,IP) are the same.

For example, let us assume that the data consists of (M=) 4 complex sinewaves and noise. Let us write down the (forward-backward) linear prediction equations $A\underline{g}=-\underline{h}$ with L=N-M/2=N-2.

$$\begin{bmatrix} y(N-2) & y(N-3) & \cdots\cdots & y(1) \\ y(N-1) & y(N-2) & \cdots\cdots & y(2) \\ y^*(2) & y^*(3) & \cdots\cdots & y^*(N-1) \\ y^*(3) & y^*(4) & \cdots\cdots & y^*(N) \end{bmatrix} \begin{bmatrix} g_1 \\ g_2 \\ . \\ . \\ g_L \end{bmatrix} = - \begin{bmatrix} y(N-1) \\ y(N) \\ y^*(1) \\ y^*(2) \end{bmatrix} \qquad 4.39$$

Since the matrix A has only (M=)4 rows, the rank of A is M (assuming that N>2M). Therefore, in this case $A=A_M$ and hence the SVD calculations to find a matrix of rank M are unnecessary. The coefficient vector $\underline{g}$ can be determined by using the

pseudoinverse of A directly.

$$\underline{g} = -A^{\#}\underline{h} = -A^{\dagger}(AA^{\dagger})^{-1}\underline{h} \qquad 4.40$$

The pseudoinverse $A^{\#}=A^{\dagger}(AA^{\dagger})^{-1}$ is used in this case (instead of $(A^{\dagger}A)^{\#}A^{\dagger}$) because M being far less than L, the matrix $AA^{\dagger}$ will be a small order matrix and can be inverted simply.

This simplification in computation goes with some reduction in performance compared to the choice of the 'optimum' value of L (in between M and N-M or N-M/2). The reason for this is that although the maximum allowed value of L is used (note that L occurs in the denominator of equation 4.33), the perturbation vector $\underline{g}$ tends to be longer. This is due to the increased noise in the principal eigenvectors of $A^{\dagger}A$ (though we do not calculate them explicitly), since the elements of $A^{\dagger}A$ are averaged over only M terms.

4.5: RELATIONSHIP TO OTHER WORKS:

Several authors in the past have used the following two basic ideas we have used in this chapter.

(1) If a vector $\underline{g}' = (g_0, g_1, \ldots\ldots g_L)^T$ is perpendicular to a set of vectors, $(1, e^{-s_k^*}, e^{-2s_k^*}, \ldots\ldots e^{-Ls_k^*})^T$, k=1,2,...M, L M, then $G(z) = \sum_{k=0}^{L} g_k z^{-k}$ has zeros at e^{s_k}, k=1,2,...M. This is Prony's idea.

(2) The M principal eigenvectors of an autocorrelation matrix ($A^\dagger A$ or $A'^\dagger A'$) are more robust to noise perturbations and they contain the relevant signal information for high enough SNR. In this section we shall describe the connection between our results and those of others.

4.5.1: THE USE OF PRINCIPAL EIGENVECTORS:

The principal eigenvectors of an autocovariance matrix have been used to advantage in other areas of research. In multivariate statistics, when a large number of measurements are available, the question often asked is whether if these could be replaced by a 'fewer number of measurements or of their functions without loss of much information, for convenience in analysis and interpretation of data'. The topic of principal component analysis regards linearly transforming the data along the first few principal eigenvectors of the covariance matrix of the data [94]. The original work in this area is due to Hotelling [85]. Rao [86,95] showed that several problems in multivariate analysis have solutions based on calculating the first few eigenvectors of a correlation matrix.

SVD has become one of the most important tools of modern numerical algebra [61]. Stable solutions to linear equations $A\underline{x}=\underline{b}$ are routinely obtained by using SVD of A and retaining only the first few singular values and the corresponding eigenvectors of A if the rest of the singular values are insignificant. Determining what is 'insignificant' is difficult and is problem dependent. SVD is also used in determining the numerical rank of a matrix [51]. Some of the original work in this area is due to Golub [84].

4.5.2: RELATED RESEARCH IN THE AREA OF SENSOR ARRAY SIGNAL PROCESSING:

Most of the work in signal processing research which involves eigenvalue-eigenvector decomposition of correlation/covariance matrices is related to sensor array signal processing. In this area, a problem is to find the angles of arrival of plane waves at a line array of equally spaced sensors.

Let us assume that we have L+1 equally spaced sensors and the waveforms received at these sensors are linear combinations of M incident plane wavefronts and noise. These waveforms (after frequency translation) are sampled at regular intervals in time. Let $\underline{y}_n$ denote the nth snapshot of the waveforms at the time instant n.

$$\underline{y}_n = (y_n(1)\ y_n(2)\ \ldots\ldots\ y_n(L+1))^T \qquad 4.41$$

The nth snapshot (see ref. [66]) at the kth sensor consists of independent noise samples $w_n(k)$, with variance σ^2, plus M source voltages.

$$y_n(k) = \sum_{i=1}^{M} a_{in} e^{jkw_i} + w_n(k),\ k=1,2,\ldots L+1 \qquad L \geq M \qquad 4.42$$

where a_{in} is the complex amplitude of the voltage induced at the i th sensor at the n th instant. $w_i = 2\Pi(q\sin(\theta_i)/\lambda)$, where θ_i is the angle of arrival of the i th plane wave, q is the spacing between the sensors and λ is the wavelength of propagation. The problem is to determine w_i's given K snapshots $\underline{y}_n$, n=1,2,...K. Note that once w_i's are determined θ_i's are also determined. If only one snapshot $\underline{y}_n$ is available for processing, then the problem reduces to our original problem described in section 1.1 (specialized to undamped sinewaves and with L+1=N). Several twists can be introduced into the above problem; for example, the noise samples $w_n(k)$ can be assumed to be correlated from sensor to sensor or the a_{in}'s can be temporally correlated or, in the limiting case, they can be sinewaves in the temporal dimension [96]. The covariance matrix R is given by

$$R = E(\underline{y}_n\ \underline{y}_n^{\dagger}) \qquad 4.43$$

where E denotes the expectation operator. Assuming that the noise samples are

uncorrelated from sensor to sensor, R can be written as (see ref. [68], for example)

$$R = \sum_{m,n=1}^{M} c_{mn} \underline{S}_m \underline{S}_n^{\dagger} + \sigma^2 I \qquad 4.44$$

where the (M×M) matrix C (with elements c_{mn}) is the covariance matrix of the source signals, $\underline{s}_k$, k=1,2,...M are the direction vectors associated with the M sources.

$$\underline{s}_k = (1, e^{jw_k}, e^{2jw_k}, \ldots\ldots e^{Ljw_k})^T \qquad 4.45$$

where w_k are related to the angles of arrival θ_k.

The methods for estimating the w_k's based on the eigenstructure of R utilizes the fact we have been using in this dissertation, i.e., if a vector $\underline{g}'$ is orthogonal to $\underline{s}_k$, (i.e., $\underline{s}_k^{\dagger} g'=0$,k=1,...M) then G(z) has zeros at e^{jw_k}, k=1,2,...M. The L+1-M eigenvectors of R (corresponding to the eigenvalue σ^2) have this property [68]. If L=M, then there is only one such eigenvector. This case coincides with Pisarenko's [60]. Several researchers [60,64-72,74,75] have used this property.

The difference between these works and ours is that we realized that the above properties are adaptable for processing N samples of a scalar time series which consists of M exponential signals (damped or undamped). That is, by forming a matrix $A'^{\dagger}A'$(as in chapter 2 or 4) where A' is a ((N-L)x(L+1) or 2(N-L)x(L+1)) data matrix using the N samples of data, we found that the eigenvectors of A' A' have similar properties to R in equation 4.44. In addition we could choose the value of L, which alters the order of $A'^{\dagger}A'$ and the degree of the polynomial G(z) to maximize the accuracy of estimation. Whereas, in this array problem, the degree L of the polynomial G(z) is dictated by the number of sensors and not by the observation time.

Usually, the matrix R in equation 4.44 is not known exactly. It is estimated from the K snapshots observed at the output of the array as

$$\hat{R} = \sum_{n=1}^{K} \underline{y}_n \underline{y}_n^{\dagger} \qquad 4.46$$

Now we shall offer some specific comments on some eigenvector based methods which are closely related to our work. We shall call the eigenvectors of R, $\lambda_1 \geq \lambda_2 \geq \ldots\ldots \geq \lambda_{L+1}$. The principal eigenvectors of R, $\underline{e}_1$, $\underline{e}_2$, ... $\underline{e}_M$ are said to span the signal subspace and $\underline{e}_{M+1}$, $\underline{e}_{M+2}$, ... $\underline{e}_{L+1}$ are said to span the noise subspace.

Pisarenko [60]: He was the first to realize the important properties of the eigenvectors of the correlation matrix R in equation 4.44. But the algorithm he proposed restricts the value of L to M and hence the accuracy of the estimates obtained using his method is poor.

Cantoni and Godara [64]: They allowed for L greater than M unlike Pisarenko. The L=M extraneous zeros of the eigenvector polynomials are eliminated by calculating the powers associated with each zero. The null space eigenvectors are not combined to form a single polynomial as done in Reddi's method [68] described below, or our IP method (see section 4.3). Hence they do not exploit the SNR improvement that stems from using all the L+1-M noise subspace eigenvectors or equivalently the M signal subspace eigenvectors.

Owsley [72,82]: He was one of the first researchers to use eigenvector methods to find source directions. But he used discrete Fourier transform processing on the M principal eigenvectors of R. Hence his was not a 'high resolution' method in that it cannot resolve closely spaced sources. He realized this in recent work [97].

Liggett [73]: He was the first researcher to use the eigenvalues of $\hat{R}$ to determine the number of signal sources. He fitted a multivariate signal model to the observed data unlike others. His wavefront model is more general than the plane wavefront model used by others.

Schmidt [67]: He has suggested a method by which a high resolution angular spectrum p(w) is computed using the L+1-M noise subspace eigenvectors. In this

method p(w) is computed as

$$p(w) = \frac{1}{\sum_{k=M+1}^{L+1} |E_k(e^{jw})|^2} \qquad 4.47$$

where $E_k(e^{jw}) = \sum_{m=0}^{L} e_{km} z^{-m} \Big|_{z=e^{jw}}$. e_{km}, m=0,1,2...L are the elements of the k th eigenvector of R. That is,

$$\underline{e}_k = [e_{k0}, e_{k1}, \cdots e_{kL}]^T \qquad 4.48$$

Since each noise subspace eigenvector polynomial will have zeros at e^{jw_k}, k=1,2,...M on the circle |z| =1 (if R was known exactly), it will produce spikes in p(w) at those angular locations. He also alludes to the frequency estimation problem in the case of a scalar time series, briefly. It is not clear how he formed the correlation matrix in this case. As we have seen in section 4.4 and in ref. [54], this has a tremendous effect on the accuracy of the frequency estimates. His method is applicable only to undamped sinewaves.

Reddi [68]: Reddi's method is closely related to ours. He uses the M principal eigenvectors $\underline{e}_1$, $\underline{e}_2$, ... $\underline{e}_M$ of $\hat{R}$ to determine a vector $\underline{g}'$ which lies perpendicular to these. Hence his method is essentially identical to our IP method (see section 4.3). He fixes the first element of $\underline{g}'$ to unity and computes the rest of the elements of $\underline{g}'$ using a pseudoinverse. Hence he implicitly uses the constraints on the polynomial G(z), g_0=1 and $|g_1|^2 + |g_2|^2 + |g_3|^2 +$ ..... + $|g_i|^2$ is minimum, though he did not realize it. This explains his observation, 'that the extraneous roots of the principal polynomial (G(z)) rarely lie on the unit circle'. This follows from our results in chapter 2.

Further comparison of these eigenvector methods can be found in our paper [98].

4.5.3: OTHER WORKS:

The following papers deal with a scalar time series (in constrast to those in the previous section) and are closely related to our work.

Henderson [57]: He computes the smaller (noise subspace) singular values and the corresponding eigenvectors of a data matrix A'. But instead of computing an L th degree polynomial G(z) as we do in section 4.3, he reduces the degree of the polynomial to M by using a Gaussian elimination procedure on the L+1-M noise subspace eigenvectors of $A'^{\dagger}A'$. Although this eliminates the nuisance caused by the L-M extraneous zeros of G(z), it results in poor parameter estimates [99].

Van Blaricum [35,100]: In ref. [35] he compared the performance of Henderson's method (above) with his method [100] (which is essentially Pisarenko's idea of using the minimum eigenvalue-eigenvector polynomial). In ref. [35] he raised relevant questions such as how to combine the L+1-M noise subspace eigenvectors of $A'^{\dagger}A'$ and how to eliminate the L-M extraneous zeros of G(z), but did not provide effective answers. We provide answers to some of his questions in section 4.3.

Holt and Antill [76]: They suggested a SVD based method to determine the number of signal components in the data using Prony's method. They had essentially suggested the same method we called the TK method in section 4.2. However, they missed a crucial point, in that they did not realize the importance of using a value of L larger than M. But the numerical example they chose has a SNR well below the threshold SNR (our numerical calculations show this). Therefore, even with an optimum choice of L (which they did not use) they would not have obtained accurate parameter estimates.

Gueguen et al. [101]: In this paper the authors present a sequence of constrained minimization problems which includes various linear prediction methods and Pisarenko method. They restate the linear prediction problem as

$$\underset{\text{w.r.t } \underline{g}'}{\text{Minimize}} \quad \underline{g}'^{\dagger} R' \, \underline{g}' \text{ with } g_0 = 1 \qquad 4.49$$

If $R' = A'^{\dagger}A'$, the above problem is identical to the forward, backward or forward-backward linear prediction methods depending on whether A' equals A'_f or A'_b or A'_{fb},

respectively, Pisarenko's method is restated as

$$\underset{\text{w.v.t } \underline{g}'}{\text{Minimize}} \quad \underline{g}'^{\dagger} R' \, \underline{g} \quad \text{with } \underline{g}'^{\dagger} \underline{g}' = 1 \qquad 4.50$$

The vector $\underline{g}'$ which minimizes the above quadratic form can easily be shown as the eigenvector corresponding to the minimum eigenvalue of R'. With respect to the above, our IP method (in section 4.3) can be rewritten as

$$\underset{\text{w.r.t } \underline{g}'}{\text{Minimize}} \quad \underline{g}'^{\dagger} R'_M \, \underline{g}' \quad \text{with } g_0 = 1 \text{ and } \phi |G(z)|^2 dz \text{ is minimum} \qquad 4.51$$

where R'_M is the rank M approximation to R'. The minimum value of the quadratic form, of course, will be zero.

The actual improvement in performance of our methods over those given above stem from properly choosing the value of L, from using the M principal eigenvector of R' (i.e. R'_M instead of R') and from choosing the appropriate type of data matrix.

CHAPTER 5: SIMULATION RESULTS:

In this chapter we quantify the performance of different algorithms we have proposed and others by measuring the bias and mean-square error of the parameter estimates $\{s_k\}$ over 500 independent trials. The mean square errors of the estimates are compared to the Cramer-Rao (CR) bound. The CR bound specifies the lower bound on the variance of (unbiased) parameter estimates. Therefore these bounds serve as a goal for our estimation algorithms. The data samples for the simulation experiments are generated in a computer as described below.

SIMULATION 1: UNDAMPED SINUSOIDS:

In this simulation 25 data samples were generated using the following formula.

$$y(n+1) = a_1 e^{j2\Pi f_1 n} + a_2 e^{j2\Pi f_2 n} + w(n), \quad n=0,1,2,\ldots 24 \qquad 5.1$$

where $f_1 = 0.52$ Hz, $f_2 = 0.5$ Hz, $a_1 = e^{j\phi 1}$, $\phi_1 = \Pi/4$, $a_2 = e^{j\phi 2}$, $\phi_2 = 0$. $w(n)$ is a computer generated, white, complex Gaussian noise sequence with variance $2\sigma^2$. σ^2 is the variance of the real and imaginary part of $w(n)$. SNR is defined as $10 \log_{10} (|a_1|^2/2\sigma^2)$. For every trial (500 such trials were used) the signal part of the data sequence $y(n)$ is fixed but $w(n)$ is different.

SIMULATION 2: EXPONENTIALLY DAMPED SINUSOIDS:

For this simulation the following formula is used in data generation.

$$y(n+1) = a_1 e^{(-\alpha_1 + j2\Pi f_1)n} + a_2 e^{(-\alpha_2 + j2\Pi f_2)n} + w(n), \quad n=0,1,2\ldots 24 \qquad 5.2$$

where

$$\alpha_1 = 0.1,\ f_1 = 0.52\text{Hz},\ a_1 = e^{j\phi 1},\ \phi_1 = 0$$

$$\alpha_2 = 0.2,\ f_2 = 0.42\text{Hz},\ a_2 = e^{j\phi 2},\ \phi_2 = 0$$

The characteristics of the noise sequence, w(n) are the same as above. The SNR is defined as 10 log $(|a_1|^2/2\sigma)^2$. Note that this is the peak SNR.

5.1: DIFFERENT METHODS USED IN THIS SIMULATION:

MAXIMUM LIKELIHOOD (ML) METHOD:

In this method the error E

$$E = \sum_{n=1}^{25} \left| y(n) - \sum_{k=1}^{2} \hat{a}_k e^{j2\pi \hat{f}_k (n-1)} \right|^2 \qquad 5.3$$

is minimized by finding $\hat{f}_1$, $\hat{f}_2$, $\hat{a}_1$ and $\hat{a}_2$ using a search procedure [18]. It is used here for estimating frequencies of undamped sinusoids only (simulation 1).

FORWARD-BACKWARD LINEAR PREDICTION (FBLP) METHOD:

The following linear equations are set up using the 25 data samples.

$$A_{fb}\, \underline{g} \cong -h_{fb} \qquad 5.4$$

Then, $\underline{g}$ is given by

$$\underline{g} = -(A_{fb}^{\dagger} A_{fb})^{-1} A_{fb}^{\dagger} \underline{h}_{fb} \qquad 5.5$$

The polynomial $G(z) = 1 + \sum_{k=1}^{L} g_k z^{-k}$ is formed and its zeros are found. The two roots closest to the unit circle are taken as the signal zeros. The angles of these zeros ($2\pi f_1$ and $2\pi f_2$) divided by two pi gives the frequency estimates. This method is applicable to undamped sinusoids only (simulation 1).

BACKWARD LINEAR PREDICTION (BLP) METHOD:

This method is meant for estimating the parameters of exponentially damped sinusoidal signals (simulation 2). In this case the linear equations used are

$$A_b\, \underline{g} \cong -\underline{h} \qquad 5.6$$

The $\underline{g}$ vector is given by

$$\underline{g} = -(A_b^{\dagger} A_b)^{-1} A_b^{\dagger} \underline{h} \qquad 5.7$$

The polynomial G(z) is formed and its roots are calculated. From the two roots that fall outside the circle $|z|=1$, the estimates of $\hat{\alpha}_1$, $\hat{f}_1$ and $\hat{\alpha}_2$ and $\hat{f}_2$ are found. A glance at figure 3.5 will help in understanding this method. If there are more than two (=M) roots outside the circle, the method fails.

MODIFIED PRONY (MP) METHOD:

In this method the linear prediction equations in equation 5.4 (for undamped sinewaves; simulation 1) or in equation 5.6 (for exponentially damped sinewaves; simulation 2) can be used. The vector $\underline{g}$ is calculated as before and the roots of G(z), $e^{\hat{s}_k}$, k=1,2,...L are found. Then we find (M=) 2 out of the L exponentials $e^{\hat{s}_k}$,k=1,2,...L that best fit the data in the sense that they give minimum value for E, defined below.

$$\underset{e^{\hat{s}_{k_i}},\ i=1,2,\ldots M}{\underset{\hat{s}_{k_i}}{\text{Min}}} \quad E = \sum_{n=0}^{24} \left| y(n+1) - \sum_{i=1}^{M} \hat{a}_{k_i} e^{\hat{s}_{k_i}} \right|^2 \qquad 5.8$$

In the case where the equation $A_b \underline{g} \cong \underline{h}_b$ is used, the zeros of G(z) that are outside the circle can be first selected. For example, assume that K out of L zeros of G(z) are outside the circle. Then we flip them inside the circle (i.e., take the mirror image). Then we find M=2 out of the K exponentials which best fit the data. Since K will be less than L the computational burden is reduced.

TUFTS-KUMARESAN (TK) METHOD:

As before, we set up the linear equations as in equation 5.4 (for undamped sinewaves) or as in equation 5.6 (for exponentially damped sinewaves). Let the following equations represent either of these two sets of linear equations.

$$A\underline{g} \cong -\underline{h} \qquad 5.9$$

We compute the SVD of A. $A = U\Sigma V^{\dagger}$. We set the smaller singular values, σ_{M+1}, σ_{M+2},..... to zero. Then we compute the vector $\underline{g}$, as

$$\underline{g} = \sum_{k=1}^{M} \sigma_k^{-1} (\underline{V}_k^{\dagger} \underline{h}) \underline{U}_k \qquad 5.10$$

where $\underline{u}_k$ and $\underline{v}_k$ are the columns of U and V. Then the polynomial G(z) is formed. The rest of the method is identical to FBLP or BLP method.

IMPROVED PISARENKO (IP) METHOD:

Using the N=25 data samples we form a matrix $A'_{fb} = (\underline{h}_{fb} \quad A_{fb})$ in the case of simulation 1 and $A'_b = (\underline{h}_b \quad A_b)$ in the case of simulation 2. Let us denote these matrices by A'. The SVD of A' is computed. $A' = P\Lambda Q$. $\underline{p}_i$ and $\underline{q}_i$ are the columns of P and Q respectively. Then the $\underline{g}$ vector is calculated as in equation 4.23. The rest of the method is identical to the TK method.

The simulation experiments are repeated at several SNR values for each method and they are tabulated below. The formulas for CR bounds for parameters of undamped sinewaves were obtained by Rife and Boorstyn [19]. We have plugged in our parameter values, SNR, and N value in their formulas to calculate the CR bounds. The formulas for CR bounds for exponentially damped sinewaves are derived in appendix C.

5.2: DISCUSSION:

The following comments regarding the simulation experiments are in order.

SIMULATION 1: UNDAMPED SINEWAVES IN NOISE: (TABLE 5.1)

(1) The ML method has the best performance as can be expected. The MSE is almost the same as the CR bound up to 3db SNR. The threshold occurs at SNR value of about 3db. By threshold, we mean that the MSE is significantly different from the CR bound at this SNR due to the occurrence of outliers.

(2) The FBLP method fails at about 25db SNR. This is because several of the L-2 extraneous zeros fall close to the unit circle at lower SNR values, resulting in several outliers. See figure 4.3.

(3) In the MP method, although the zeros obtained by rooting the polynomial G(z) are exactly the same as in FBLP method, the threshold occurs at a much lower

$\sqrt{\text{MEAN SQUARE FREQUENCY ESTIMATION ERROR}}$ OF f_1

SNR	MP METHOD	TK METHOD	IP METHOD	ML METHOD	FBLP METHOD	CR BOUND
db	L=12	L=18	L=18		L=12	
30	$0.427\text{x}10^{-3}$	$0.345\text{x}10^{-3}$	$0.356\text{x}10^{-3}$	$0.306\text{x}10^{-3}$	$0.427\text{x}10^{-3}$	$0.311\text{x}10^{-3}$
20	$0.130\text{x}10^{-2}$	$0.105\text{x}10^{-2}$	$0.114\text{x}10^{-2}$	$0.978\text{x}10^{-3}$	**	$0.984\text{x}10^{-3}$
15	$0.238\text{x}10^{-2}$	$0.184\text{x}10^{-2}$	$0.206\text{x}10^{-2}$	$0.180\text{x}10^{-2}$	**	$0.175\text{x}10^{-2}$
12	$0.373\text{x}10^{-2}$	$1.279\text{x}10^{-2}$	$0.287\text{x}10^{-2}$	$0.280\text{x}10^{-2}$	**	$0.276\text{x}10^{-2}$
10	$0.417\text{x}10^{-2}$	$0.320\text{x}10^{-2}$	$0.347\text{x}10^{-2}$	$0.344\text{x}10^{-2}$	**	$0.311\text{x}10^{-2}$
7	$0.601\text{x}10^{-2}$	$0.450\text{x}10^{-2}$	$0.456\text{x}10^{-2}$	$0.481\text{x}10^{-2}$	**	$0.490.10^{-2}$
5	**	**	**	$0.560\text{x}10^{-2}$	**	$0.553\text{x}10^{-2}$
3	**	**	**	$0.653\text{x}10^{-2}$	**	$0.696\text{x}10^{-2}$
0	**	**	**	**	**	$0.984\text{x}10^{-2}$

TABLE 5.1: Comparison of the frequency estimation error of different methods to the CR Bound. ** denotes that the mean square error is much larger than the CR bound due to the occurrence of outliers. Similar numbers are obtained for frequency f_2 as well. 500 data sets generated in eqn. 5.1 are used. The chosen L values are best for that method. At 7db SNR, for the IP, MP, and TK methods the bias is about a third of the $\sqrt{\text{mean square error}}$.

SNR because the signal zeros of G(z) are picked based on the criterion in equation 5.8, unlike as in the FBLP method. The value of MSE at 30db SNR in the case of the MP method and the FBLP method are the same because the same zeros of G(z) are picked in each trial in both methods.

(4) The performance of the TK and IP methods are about the same and are only slightly poorer than ML method. Their performance is slightly better than the MP method.

SIMULATION 2: EXPONENTIALLY DAMPED SINEWAVES IN NOISE: (TABLES 5.2-5.5)

(1) The BLP method did not give the correct estimates even at SNR=40db because there were too many extraneous zeros outside the unit circle.

(2) As in Simulation 1, the TK and IP methods perform slightly (in some cases much) better than the MP method.

(3) The bias in the estimates of the damping factors α_1 and α_2 becomes significant at lower SNR values. This bias can be reduced by using the noise singular values of A_b or A_b' in the IP and TK methods as suggested by us in ref. [55].

The results in Tables 5.1-5.5 are drawn in a graphical form in figures 5.1-5.5. Further simulation results can be found in our papers [18,50,52-56,92,98,99, 104].

SNR	TK METHOD		IP METHOD		MP METHOD		CR BOUND
db	L=18		L=18		L=12		
	bias	mse	bias	mse	bias	mse	
40	$-.279\times10^{-4}$	$.101\times10^{-2}$	$.369\times10^{-4}$	$.110\times10^{-2}$	$-.132\times10^{-4}$	$.150\times10^{-2}$	$.828\times10^{-3}$
30	$.211\times10^{-4}$	$.319.10^{-2}$	$.215\times10^{-3}$	$.347\times10^{-2}$	$-.941\times10^{-4}$	$.509\times10^{-2}$	$.262\times10^{-2}$
20	$.110\times10^{-2}$	$.100\times10^{01}$	$.164\times10^{-2}$	$.112\times10^{-1}$	$.106\times10^{-2}$	$.172\times10^{-1}$	$.828\times10^{-2}$
15	$.377\times10^{-2}$	$.176\times10^{-1}$	$.470\times10^{-2}$	$.201\times10^{-1}$	$.367\times10^{-2}$	$.297\times10^{-1}$	$.147\times10^{-1}$
12	**	**	$.891\times10^{-2}$	$.303\times10^{-1}$	$.676\times10^{-2}$	$.432\times10^{-1}$	$.219\times10^{-1}$
10	**	**	**	**	**	**	$.262\times10^{-1}$

TABLE 5.2: Estimation error for α_1 (=0.1). ** denotes that the estimation error is large compared to the CR bound.

SNR	TK METHOD		IP METHOD		MP METHOD		CR BOUND
db	L=18		L=18		L=12		
	bias	mse	bias	mse	bias	mse	
40	$-.497\times10^{-4}$	$.218\times10^{-2}$	$-.856\times10^{-5}$	$.271\times10^{-2}$	$-.207\times10^{-3}$	$.463\times10^{-2}$	$.202\times10^{-2}$
30	$-.900\times10^{-3}$	$.686\times10^{-2}$	$-.506\times10^{-3}$	$.857\times10^{-2}$	$-.644\times10^{-3}$	$.134\times10^{-1}$	$.637\times10^{-2}$
20	$-.103\times10^{-1}$	$.229\times10^{-1}$	$-.643\times10^{-2}$	$.268\times10^{-1}$	$-.216\times10^{-2}$	$.397\times10^{-1}$	$.201\times10^{-1}$
15	$-.344\times10^{-1}$	$.470\times10^{-1}$	$-.216\times10^{-1}$	$.467\times10^{-1}$	$-.191\times10^{-2}$	$.711\times10^{-1}$	$.358\times10^{-1}$
12	**	**	$-.455\times10^{-1}$	$.890\times10^{-1}$	$-.935\times10^{-2}$	$.998\times10^{-1}$	$.540\times10^{-1}$
10	**	**	**	**	**	**	$.637\times10^{-1}$

TABLE 5.3: Estimation error for α_2 (=0.2). ** denotes that the estimation error is large compared to the CR bound.

SNR	TK METHOD		IP METHOD		MP METHOD		CR BOUND
db	L=18		L=18		L=12		
	bias	mse	bias	mse	bias	mse	
40	$.370\times10^{-5}$	$.162.10^{-3}$	$-.376\times10^{-5}$	$.155\times10^{-3}$	$.400\times10^{-5}$	$.274\times10^{-3}$	$.132\times10^{-3}$
30	$.260\times10^{-4}$	$.511\times10^{-3}$	$.132.1^{-3}$	$.488\times10^{-3}$	$.398\times10^{-4}$	$.866\times10^{-3}$	$.417\times10^{-3}$
20	$.225\times10^{-3}$	$.160\times10^{-2}$	$.138\times10^{-3}$	$.151\times10^{-2}$	$.169\times10^{-3}$	$.286\times10^{-2}$	$.132\times10^{-2}$
15	$.674.10^{-3}$	$.277\times10^{-2}$	$.513\times10^{-3}$	$.262\times10^{-2}$	$.572\times10^{-3}$	$.522\times10^{-2}$	$.234\times10^{-2}$
12	**	**	$.977\times10^{-3}$	$.358\times10^{-2}$	$.150\times10^{-2}$	$.852\times10^{-2}$	$.350\times10^{-2}$
10	**	**	**	**	**	**	$.417\times10^{-2}$

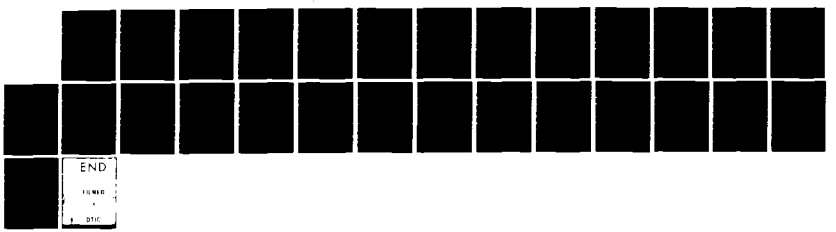
TABLE 5.4: Estimation error for f_1 (=0.52 Hz).

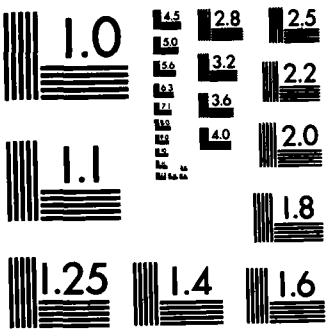
SNR	TK METHOD		IP METHOD		MP METHOD		CR BOUND
db	L=18		L=18		L=12		
	bias	mse	bias	mse	bias	mse	
40	$-.115\times10^{-4}$	$.362\times10^{-3}$	$.267\times10^{-4}$	$.350\times10^{-3}$	$.104\times10^{-3}$	$.880\times10^{-3}$	$.321\times10^{-3}$
30	$-.627\times10^{-4}$	$.114\times10^{-2}$	$.634\times10^{-4}$	$.111\times10^{-2}$	$-.253\times10^{-3}$	$.278\times10^{-2}$	$.101\times10^{-2}$
20	$-.458\times10^{-3}$	$.358\times10^{-2}$	$-.978.10^{-5}$	$.351\times10^{-2}$	$-.753\times10^{-3}$	$.778\times10^{-2}$	$.321\times10^{-2}$
15	$-.132\times10^{-2}$	$.630\times10^{-2}$	$-.439\times10^{-3}$	$.631\times10^{-2}$	$-.211\times10^{-3}$	$.166\times10^{-1}$	$.570\times10^{-2}$
12	**	**	$-.926\times10^{-3}$	$.920\times10^{-2}$	$-.231\times10^{-3}$	$.248\times10^{-1}$	$.840\times10^{-2}$
10	**	**	**	**	**	**	$.101\times10^{-1}$

TABLE 5.5: Estimation error in f_2 (=0.5Hz).

AD-A123 177 ESTIMATING THE PARAMETERS OF EXPONENTIALLY DAMPED OR UNDAMPED SINUSOIDAL.. (U) RHODE ISLAND UNIV KINGSTON DEPT OF ELECTRICAL ENGINEERING R KUMARESAN OCT 82 2/2

UNCLASSIFIED N00014-81-K-0144 F/G 17/2 NL





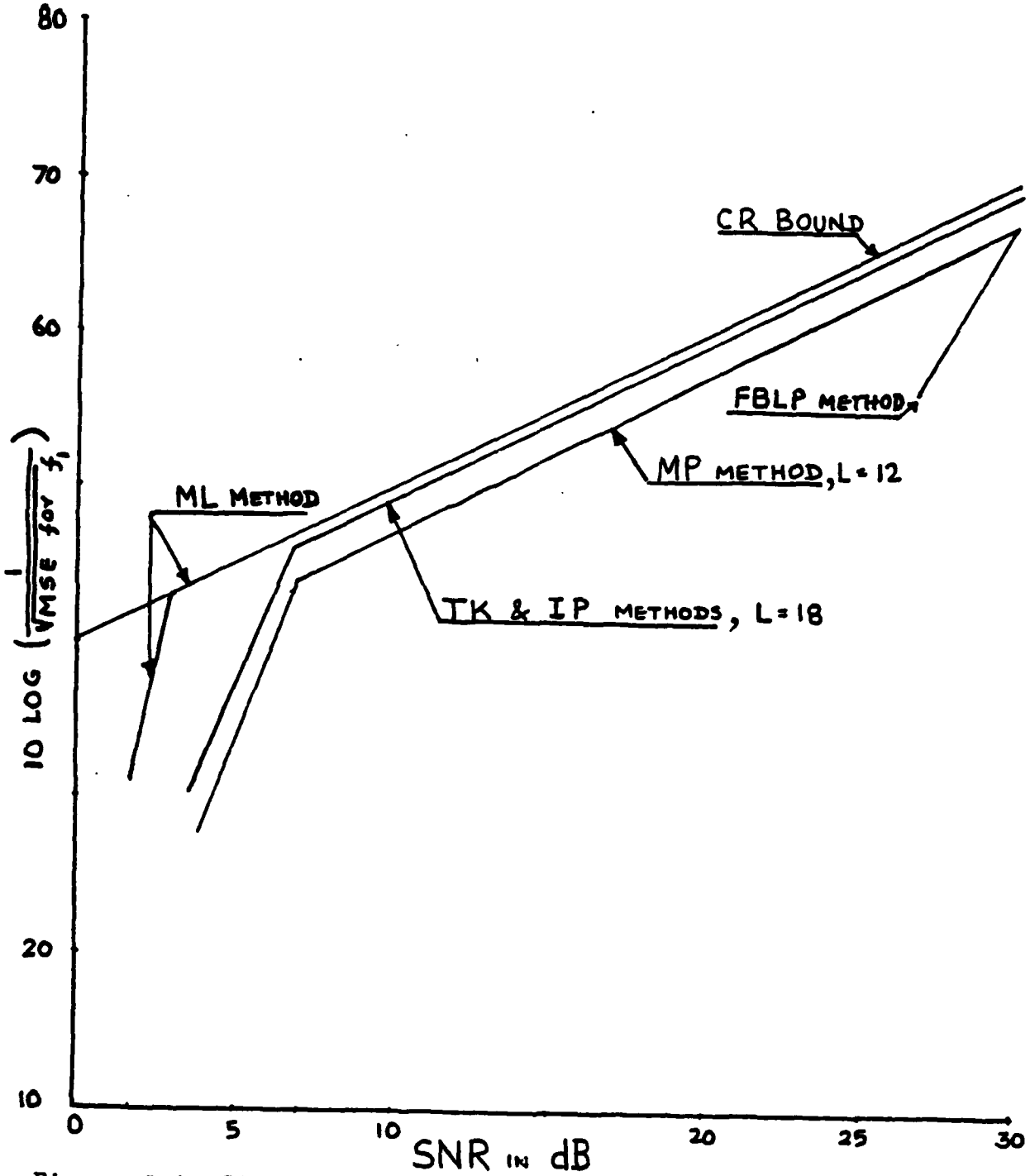


Figure 5.1: Simulation I; Performance Comparison of Different Methods; See Table 5.1.

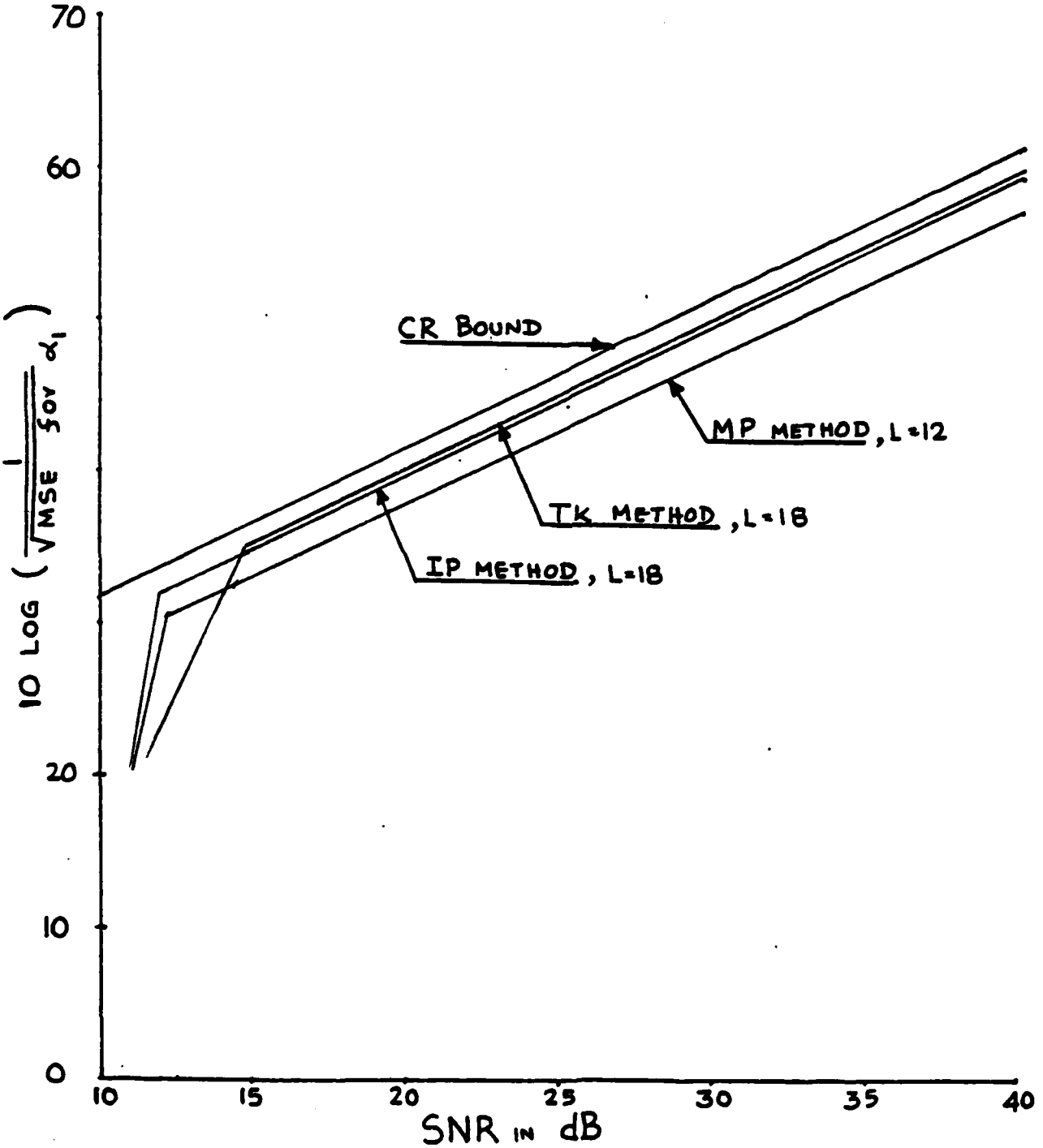


Figure 5.2: Simulation II; Performance Comparison of Different Methods; See Table 5.2.

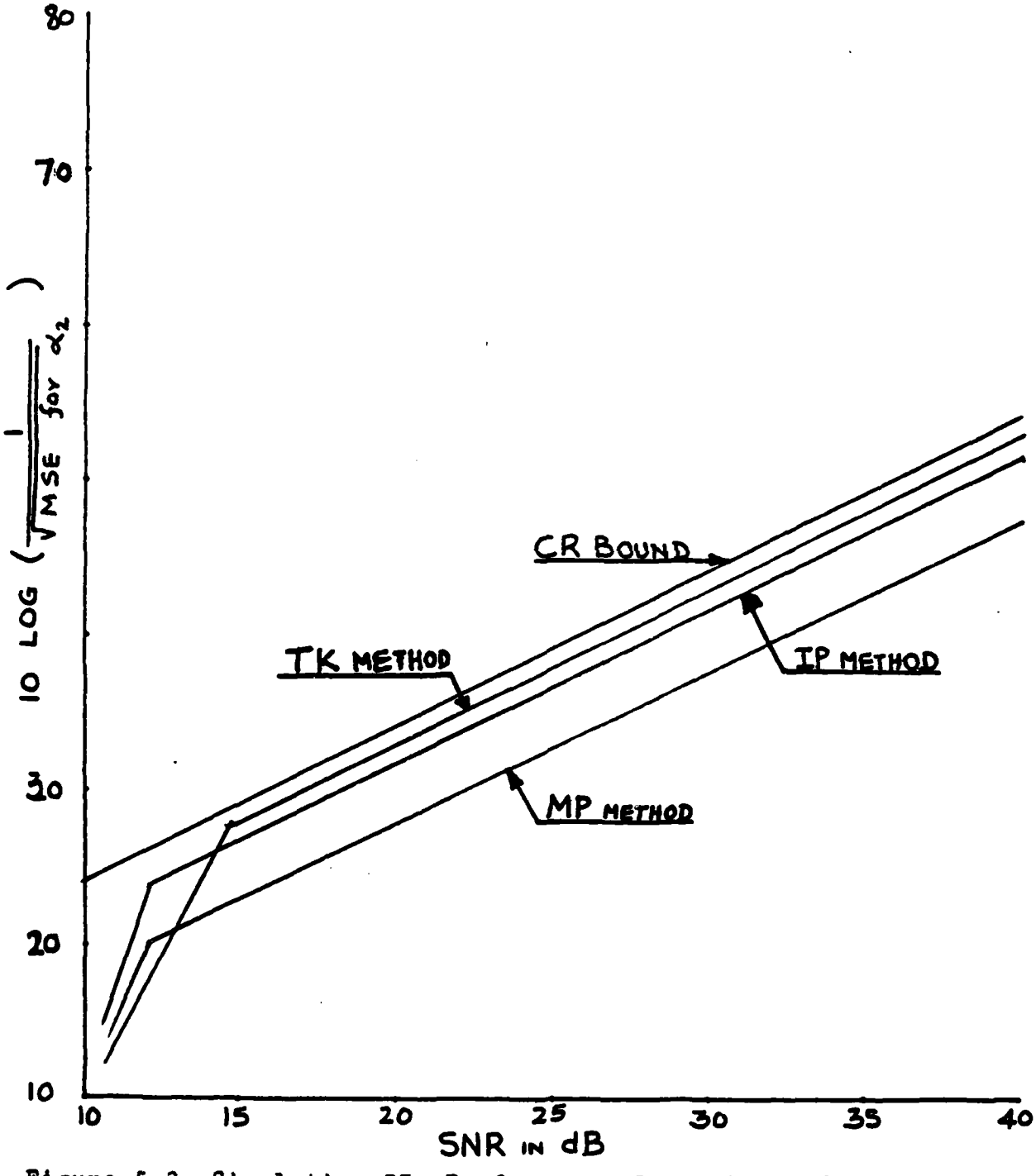


Figure 5.3: Simulation II; Performance Comparison of Different Methods; See Table 5.3.

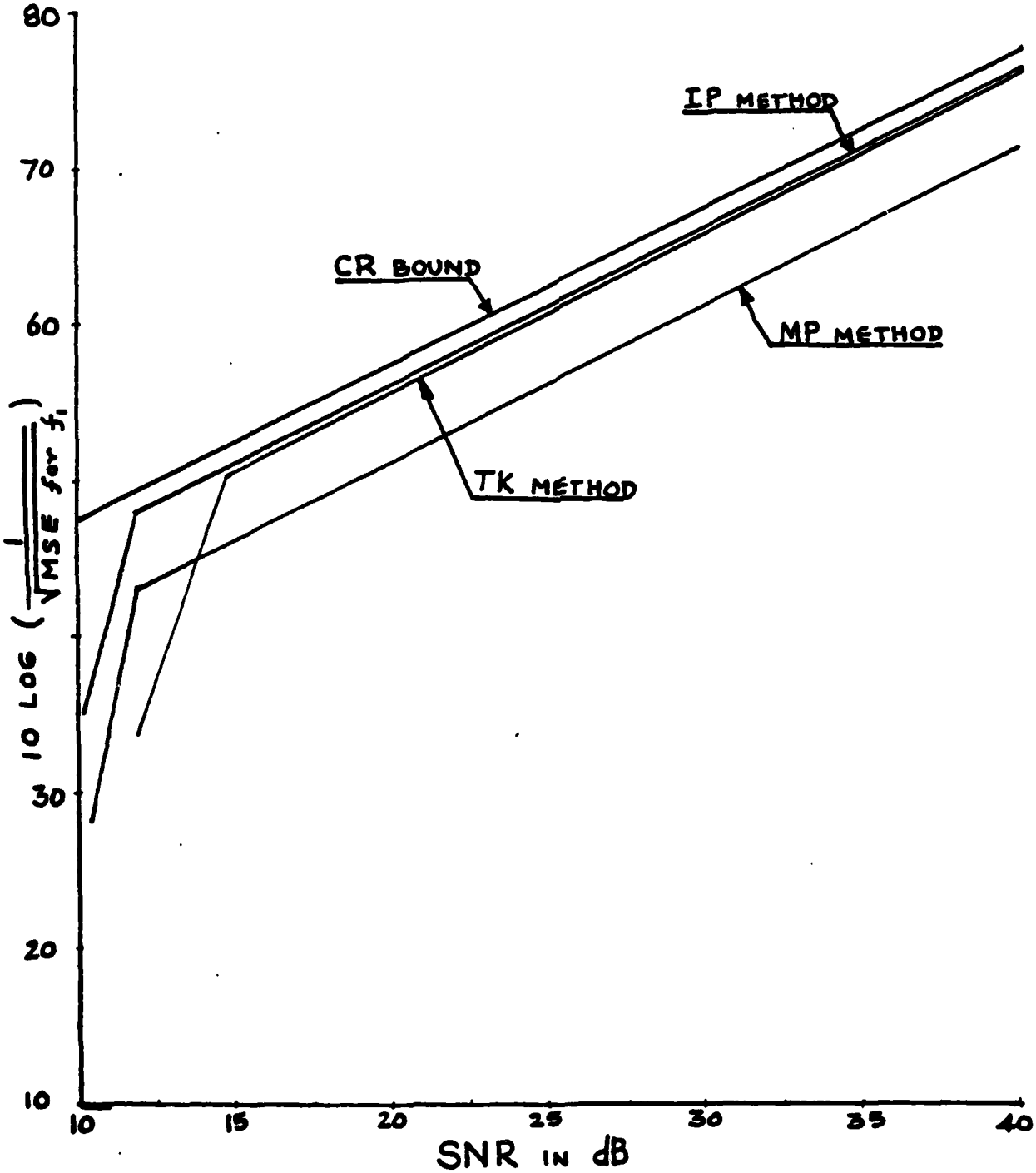


Figure 5.4: Simulation II; Performance Comparison of Different Methods; See Table 5.4.

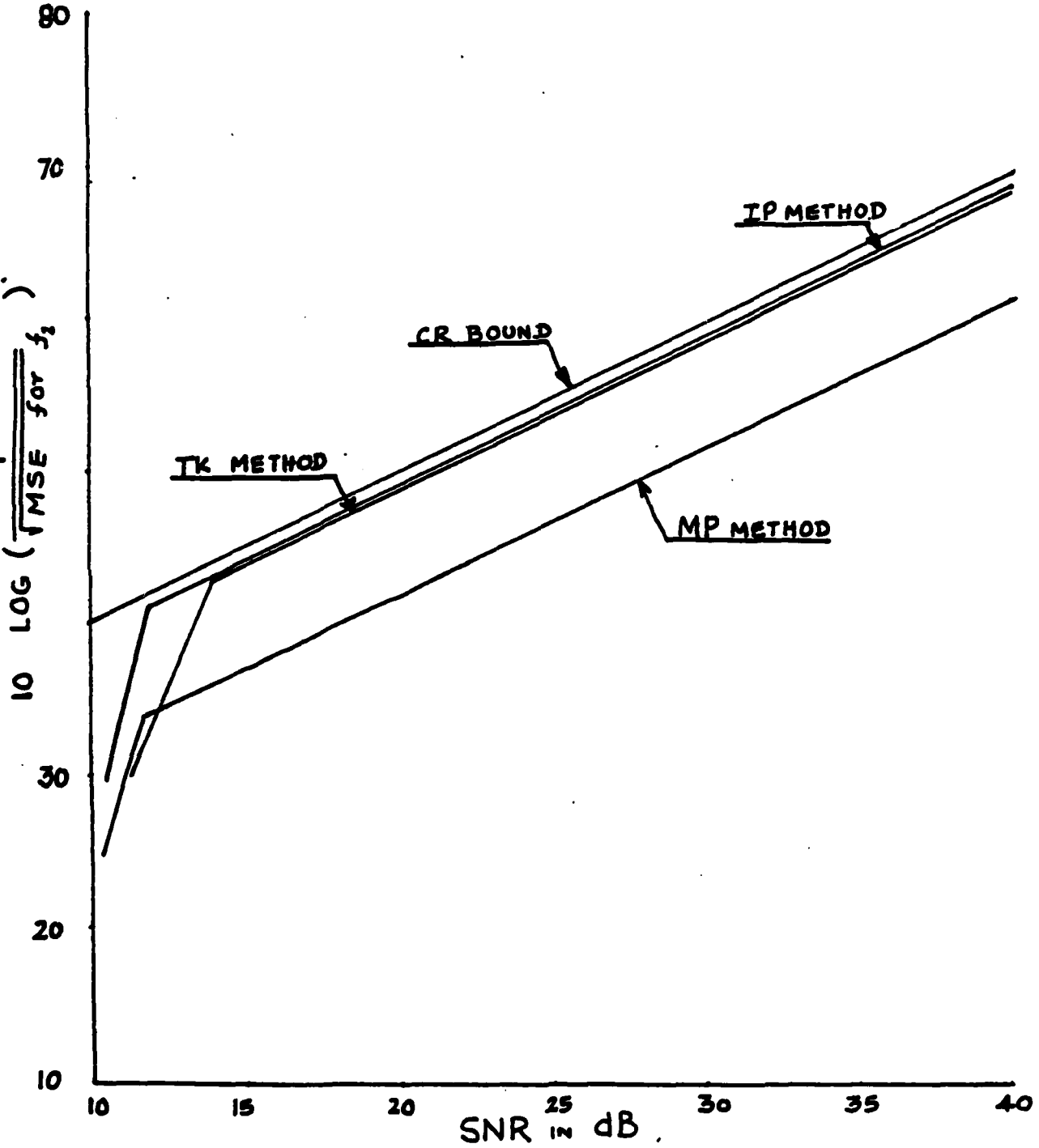


Figure 5.5: Simulation II; Performance Comparison of Different Methods; See Table 5.5.

CHAPTER 6: CONCLUSIONS

We assumed that a short record of a data sequence y(n), composed of M (unknown) number of exponentially damped or undamped sinusoidal signals in noise, i.e., $y(n) = \sum_{k=1}^{M} a_k e^{s_k n} + w(n)$, n=1,2,...N, is available for processing. We set out to estimate the values of $\{a_k\}$ and $\{s_k\}$ and M. The goal was to estimate these parameters at lower SNR values than possible using presently available algorithms. We discovered (or rediscovered, in some cases) the following main results.

(1) Estimating the parameters $\{s_k\}$ from the zeros of a polynomial G(z) (i.e., using the indirect approach originated by Prony) is preferable computationally and otherwise, especially when multiple signal components are present in the data. However, Prony's idea needs modifications to accommodate the presence of noise in the data.

(2) Selecting a value for L, the degree of G(z), greater than M improves the estimation accuracy considerably. But this alone is not sufficient for accurate parameter estimation. Because the higher degree polynomial, G(z), has L-M extraneous zeros which results in outliers.

(3) We suggested two methods for eliminating the extraneous zeros of G(z). This is important in extending the threshold SNR of the methods presented.

(3a) Firstly, we presented a selection procedure (in chapter 3) which picks the M signal zeros out of the L zeros of G(z) based on a least squares criterion. This procedure extends the usefulness of Prony's method to low-SNR values. It also gives a method of estimating the value of M.

(3b) Further improvement in accuracy of the estimates was achieved by improving the SNR in the data using SVD.

(4) The majority of the improvement in accuracy is achieved by using a value of L greater than M, which intuitively means explaining part of the noise in the data by L-M additional exponentials, paving the way for more accurate parameter estimates. This is the primary reason for extension of the threshold SNR value as shown by the results of modified Prony (MP) method. Using SVD of the data matrix further improves the accuracy.

(5) We pointed out several nice properties of the polynomial G(z), when its coefficients are constrained such that $g_0=1$ and $|g_1|^2 + |g_2|^2 + \ldots\ldots + |g_L|^2$ is minimum. These were first used in ref. [92].

(6) Other noteworthy points include (a) use of backward prediction to identify the signal zeros from the extraneous zeros (in the case of exponentially damped sinusoids), (b) using the extreme value of L (=N-M or N-M/2) to avoid expensive SVD calculations [92] (see section 4.4.2), but at the cost of some performance.

APPENDIX A: MINIMUM PHASE PROPERTY OF THE 'PREDICTION-ERROR' FILTER

In equation 2.11, the prediction-error filter polynomial $G(z)=1+\sum_{k=1}^{L} g_k z^{-k}$ is factored into two polynomials as below.

$$G(z) = G_1(z)G_2(z) \qquad A.1$$

where

$$G_1(z)=1+\sum_{k=1}^{M} b_k z^{-k} \qquad A.2$$

$$G_2(z)=1+\sum_{k=1}^{L-M} c_k z^{-k} \qquad A.3$$

The factor $G_1(z)$ has the M signal zeros located at e^{s_k} (or $e^{-s_k^*}$), k=1,2,...M as shown in section 2.1. If we find the coefficients of G(z), g_1, g_2, ... g_L such that $|g_1|^2 + |g_2|^2 + ... + |g_L|^2$ is minimum (as in proposition 5, section 2.2), we wish to show that the L-M zeros of $G_2(z)$ will lie within the unit circle. A polynomial is said to be minimum phase if all its zeros are within the unit circle.

Note that minimizing $\sum_{k=1}^{L} |g_k|^2$, is the same as minimizing the following quantity, Q,

$$Q = 1+|g_1|^2 + |g_2|^2 + ... + |g_L|^2 \qquad A.4$$

which is,

$$\begin{aligned} Q &= \int_{-\Pi}^{\Pi} |G(e^{jw})|^2 dw \\ &= \int_{-\Pi}^{\Pi} |G_1(e^{jw})|^2 \; |G_2(e^{jw})|^2 dw \end{aligned} \qquad A.5$$

We now show that the roots of $G_2(z)$ are inside the unit circle if Q is minimized, following a proof given by Lang and McClellan [87] and Pakula and Kay [88]. Let us factor $G_2(z)$ as $G_2(z) = (1-z_1 z^{-1})G_2'(z)$, where $G_2'(z) = \prod_{k=1,k\neq i}^{L-M} (1-z_k z^{-1})$ and z_i is any zero of $G_2(z)$. Let $z_i=re^{jw_i}$. Then

$$Q(r) = \int_{\Pi}^{\Pi} |G_1(e^{jw})|^2 \; |(1-re^{j(w_i-w)})|^2 \; |G_2'(e^{jw})|^2 \, dw$$

$$= \int_{\Pi}^{\Pi} |G_1(e^{jw})|^2 \ (1-2r\cos(w_i-w)+r^9) \ |G_2'(e^{jw})|^2 dw \qquad A.6$$

Differentiating Q(r) with respect to r, we get

$$\frac{dQ(r)}{dr} = 2\int_{\Pi}^{\Pi} |G_1(e^{jw})|^2 \ (2r-2\cos(w_i-w)) \ |G_2'(e^{jw})|^2 dw \qquad A.7$$

and since, for $r \geq 1$, $\frac{dQ(r)}{dr} > 0$, Q(r) cannot be minimized if a zero of $G_2(z)$ is on or outside the unit circle. Hence, $G_2(z)$ is minimum phase and the L-M extraneous zeros of G(z) lie within the unit circle.

APPENDIX B: THE EIGENVECTORS OF $A^{\dagger}A$ AND $A\,A^{\dagger}$

In this appendix, we shall show that the M principal eigenvectors of $A^{\dagger}A$ and $A\,A^{\dagger}$ span the 'signal subspace'. The term 'signal subspace' is explained below. Similar properties are true for the matrices $A'^{\dagger}A'$ and $A'\,A'^{\dagger}$ as well. The noiseless signal sequence y(n) is, $y(n) = \sum_{k=1}^{M} a_k e^{s_k n}$. Let A be the forward-data matrix, and let L satisfy the inequality, $M \leq L \leq (N-M)$.

$$A = \begin{bmatrix} y(L) & y(L-1) & \ldots\ldots & y(1) \\ y(L+1) & y(L) & \ldots\ldots & y(2) \\ \cdot & \cdot & \ldots\ldots & \cdot \\ \cdot & \cdot & \ldots\ldots & \cdot \\ y(N-1) & y(N-2) & \ldots\ldots & y(N-L) \end{bmatrix} \qquad \text{B.1}$$

After substituting for y(n), the matrix A can be decomposed into two matrices as follows.

$$A = \begin{bmatrix} a_1 e^{s_1 L} & a_2 e^{s_2 L} & \ldots\ldots & a_M e^{s_M L} \\ a_1 e^{s_1 (L+1)} & \ldots\ldots\ldots & & a_M e^{s_M (L+1)} \\ \cdot & \ldots\ldots\ldots & & \\ \cdot & \ldots\ldots\ldots & & \\ \cdot & \ldots\ldots\ldots & & \\ a_1 e^{s_1 (N-1)} & \ldots\ldots\ldots & & a_M e^{s_M (N-1)} \end{bmatrix} \begin{bmatrix} 1 & e^{-s_1} & e^{-2s_1} & \ldots\ldots & e^{-(L-1)s_1} \\ 1 & e^{-s_2} & e^{-2s_2} & \ldots\ldots & e^{-(L-1)s_2} \\ \cdot & \cdot & \cdot & \ldots\ldots & \\ 1 & e^{-s_M} & e^{-2s_M} & \ldots\ldots & e^{-(L-1)s_M} \end{bmatrix} \qquad \text{B.2}$$

Let us call the two matrices on the right side of B.2, B and C. Hence,

$$A = \underset{(N-L)xM}{B} \quad \underset{MxL}{C} \qquad \text{B.3}$$

and,

$$A^{\dagger}A = C^{\dagger}B^{\dagger}B\,C \qquad \text{B.4}$$

Now, consider the (MxM) matrix $B^{\dagger}BCC^{\dagger}$. It is positive definite because both B and C have full rank. Let us call its eigenvalues λ_i, i-1,2,...M and its

eigenvectors $\underline{w}_i = (w_{i1}, w_{12}, \ldots w_{1M})^T$, i-1,2,...M. By definition

$$B^\dagger BCC^\dagger \underline{w}_i = \lambda_i \underline{w}_i \ , \qquad i=1,2,\ldots M \qquad \text{B.5}$$

Premultiplying the above equation by $C^\dagger$, we get

$$C^\dagger B^\dagger BCC^\dagger \underline{w}_i = \lambda_i C^\dagger \underline{w}_i \ , \ i=1,2,\ldots M \qquad \text{B16}$$

Let us define $\underline{v}_i = C^\dagger \underline{w}_i$. But $A^\dagger A = C^\dagger B^\dagger BC$. Hence B.6 simplifies to

$$A^\dagger A \underline{v}_i = \lambda_i \underline{v}_i \qquad i=1,2,\ldots M \qquad \text{B.7}$$

Therefore, λ_i, i=1,2,...M are also the non-zero eigenvalues of th TxL) matrix $A^\dagger A$. The corresponding eigenvectors of $A^\dagger A$ are

$$\underline{v}_i = \begin{bmatrix} 1 & 1 & \ldots\ldots\ldots & 1 \\ e^{-s_1^*} & e^{-s_2^*} & \ldots\ldots\ldots & e^{-s_M^*} \\ \cdot & \cdot & \ldots\ldots\ldots & \cdot \\ \cdot & \cdot & \ldots\ldots\ldots & \cdot \\ \cdot & \cdot & \ldots\ldots\ldots & \cdot \\ \cdot & \cdot & \ldots\ldots\ldots & \cdot \\ e^{-s_1^* L} & e^{-s_2^* L} & \ldots\ldots\ldots & e^{-s_M^* L} \end{bmatrix} \begin{bmatrix} w_{i1} \\ w_{12} \\ \cdot \\ \cdot \\ w_{iM} \end{bmatrix} \qquad \text{B.8}$$

Thus any principal eigenvector of $A^\dagger A$ will be a linear combination of the columns of $C^\dagger$. Similarly it can be shown that any principal eigenvector of $AA^\dagger$ is a linear combination of the columns of B (and hence linear combination of vectors, $(1, e^{s_i}, e^{2s_i}, \ldots e^{(N-L)s_i})^T$, i=1,2,...M). Any N-L contiguous samples of the signal sequence (for example, $(y(k), y(k+1), \ldots . y(N-L+k-1))^T$ can be expressed as a linear combination of the M principal eigenvectors of $AA^\dagger$. Hence these vectors are said to span the 'signal subspace'.

Extending the above, expressions for the eigenvalues and the signal subspace eigenvectors of $AA^\dagger$ and $A^\dagger A$ can be derived as shown in ref. [89]. Similar expressions can also be derived for the backward data matrix A_b and forward-backward data matrix A_{fb}.

APPENDIX C: CRAMER-RAO BOUND FORMULAS FOR THE PARAMETERS OF EXPONENTIALLY DAMPED SINUSOIDAL SIGNALS:

In this appendix we shall describe an outline of the calculation of the CR bounds for the parameters of one (M=1) and two (M=2) exponentially damped signals in white Gaussian noise. At high SNR values our parameter estimates are essentially unbiased. The CR bounds, in that case, provide a lower bound on the variance of the parameter estimates. The data samples y(n) are given by the formula

$$y(n) = \sum_{k=1}^{M} b_k e^{j\phi_k} e^{-\alpha_k n^{\dagger} j2\Pi f_k n} + w(n), \quad n=0,1,\ldots N-1 \qquad C.1$$

where w(n) is a white Gaussian sequence. The sampling interval is assumed one second. The real and imaginary parts of w(n) have a variance σ^2. Comparing equation 1.1 with the above formula, we see that $a_k=b_k e^{j\phi k}$ and $s_k=-\alpha_k+j2\Pi f_k$, k=1,2,...M. For convenience we shall relabel the 4M parameters as follows.

$$\begin{aligned} \beta_{4(k-1)} &= f_k \\ \beta_{4(k-1)+1} &= c_k \\ \beta_{4(k-1)+2} &= \phi_k \\ \beta_{4(k-1)+3} &= \alpha_k \end{aligned} \qquad k=1,2,\ldots M \qquad C.2$$

$$\underline{y}=(y(0),y(1),\ldots\ldots\ldots\ldots\ldots y(N-1))^T \qquad C.3$$

$$\beta= (\beta_0, \beta_1, \beta_2,\ldots\ldots\ldots\ldots\beta_{4(M-1)+3})^T \qquad C.4$$

The probability density function for the data vector conditioned on the unknown parameter vector can be written as

$$P(\underline{y}/\underline{\beta}) = (2\Pi\sigma^2)^{-N} \exp \frac{-1}{2^2} \sum_{n=0}^{N-1} \left| y(n) - \sum_{k=1}^{M} b_k e^{j\phi_k} e^{(-\alpha_k+j2\Pi f_k)n} \right|^2 \qquad C.5$$

The CR bound states [93] that for any unbiased estimate of $\underline{\beta}$

$$\text{Var}\ (\hat{\beta}_i) \geq [J^{-1}(\underline{\beta})]_{ii} \qquad C.6$$

where $J(\underline{\beta})$ is the (4Mx4M) Fisher information matrix with elements [93]

$$[J(\underline{\beta})]_{ij} = -E \left[\frac{\partial^2}{\partial\beta_i \partial\beta_j} \ln P(\underline{y}/\underline{\beta}) \right]_{\underline{\beta}} \quad \text{C.7}$$

We shall consider the cases M=1 and 2. The derivaton follows that of Rife and Boorstyn [19] who derived similar expressions for the case of sinusoidal signals.

CASE 1: M=1

Using formula C.7 above we can write down the Fisher information matrix for this case as follows

$$J = J_{ii} = \begin{bmatrix} b_i^2 P_{i2} & 0 & b_i^2 P_{i1} & 0 \\ & P_{i0} & 0 & -b_i P_{i1} \\ & \text{symmetric} & c_i^2 P_{i0} & 0 \\ & & & c_i^2 P_{i2} \end{bmatrix}, \quad i=1 \quad \text{C.8}$$

where

$$P_{ij} = \sum_{n=0}^{N-1} n^i e^{-2\alpha_i n}, \quad j=0,1,2 \quad \text{C.9}$$

Inverting J analytically, we find the diagonal terms of J^{-1} which provide the bounds on the variance for the parameters f_1 and α_1 as follows.

$$\text{Var}(f_1) \geq \frac{\sigma^2}{4\Pi^2 b_1^2} \quad \frac{P_{10}}{P_{10}P_{12} - P_{11}^2} \quad \text{C.10}$$

$$\text{Var}(\alpha_1) \geq \frac{\sigma^2}{b_1^2} \quad \frac{P_{10}}{P_{10}P_{12} - P_{11}^2} \quad \text{C.11}$$

CASE 2: M=2

In this case J is an (8x8) matrix. We partition J as follows

$$J = \left[\begin{array}{c:c} J_{11} & J_{12} \\ \hdashline J_{12}^T & J_{22} \end{array} \right] \quad \text{C.12}$$

where the matrices J_{11} and J_{22} are given by equation C.8 and C.9 with i=1 and i=2 respectively. J_{12} is a matrix exhibiting the interaction between the

parameters of the two signals.

$$J_{12} = \begin{bmatrix} b_1b_2r_2 & b_1q_1 & b_1b_2r_1 & -b_1b_2q_2 \\ -b_2q_1 & r_0 & q_0 & -b_2r_1 \\ b_1b_2r_1 & b_1q_0 & b_1b_2r_0 & -b_1b_2q_1 \\ b_1b_2q_2 & -b_1r_1 & b_1b_2q_1 & b_1b_2r_2 \end{bmatrix} \qquad \text{C.13}$$

where

$$\left.\begin{aligned} r_j &= \sum_{n=0}^{N-1} e^{-(\alpha_1+\alpha_2)n} n_{nj} \cos\Delta_n \\ q_j &= \sum_{0\,n=0}^{N-1} e^{-(\alpha_1+\alpha_2)n} n_{nj} \sin\Delta_n \end{aligned}\right\} \quad j=0,1,2 \qquad \text{C.14}$$

$$\Delta_n = 2\Pi(f_2 - f_1)n + \theta_2 = \theta_1$$

The matrix J was inverted using a machine to compute the CR bounds of the parameters f_1, f_2, α_1, and α_2 which are used in chapter 5. Interestingly, due to the special structure of J, it turns out that the diagonal elements of J^{-1} are independent of $\phi_2-\phi_1$.

REFERENCES:

1) C.G.M.Fant, Acoustic Theory of Speech Production, Mouton and Company, The Hague, 1960.

2) E.N.Pinson, "Pitch Synchronous Time-Domain Estimation of Formant Frequencies and Bandwidths", Journal of Acoust. Soc. of Amer., Vol. 35, pp.1264-1273, August 1963.

3) S.M.Kay and L.Marple, "Spectrum Analysis-A Modern Perspective", Proceedings of the IEEE, Vol.69, No.11, pp.1380-1419, November 1981.

4) R.N.Bracewell, "Radio Interferometry of Discrete Sources", Proc. of IRE, Vol.46, pp.97-105, January 1958.

5) K.Davies, "Ionospheric Radio Propagation", Nat. Bur. Stand. Monogr.,80,191, 1965.

6) J.Myhill et al. "Investigation of an operator method in the Analysis of Biological Tracer Data", Biophysical Journal, Vol.5, pp.89-107,1965.

7) W.J.Wiscombe and J.W.Evans, "Exponential Sum Fitting of Radiative Transmission Function", Jour. of Computational Physics, Vol. 24, pp.416-444, 1977.

8) R.K.Mehra, "On-Line Identification of a Linear Dynamical System with Applications to Kalman-Filtering", IEEE Trans. on Auto. Control, AC-16, P.12, Feb.1971.

9) A.A.Beex and L.L.Scharf, "Covariance Sequence Approximation for Parametric Spectrum Modelling", IEEE Trans. on Acoust. Speech and Signal Proc., Vol. 29, p. 1042, Oct. 1981.

10) A.Schuster, "On the Investigation of Hidden Periodicities with Application of a supposed 26 day period of Meteorological phenomenon", Terrestrial Magnetism, Vol. 3, pp. 13-41, March 1898.

11) A.Papoulis, "A New Algorithm in Spectral Analysis and Band Limited Signal Extrapolation", IEEE Trans. on Circuits and Systems, Vol. CAS-22, pp. 735-742, Sept. 1975.

12) R.W.Gerchberg, "Superresolution Through Error Energy Reduction", Opt. Acta, Vol. 21, No. 9, pp. 709-720, 1974.

13) K.J.Astorm and P.Eykhoff, "System Identification-A Survey", Automatica, Vol. 7, pp. 123-162, 1971.

14) D.W.Marquardt, "An Algorithm for Least Squares Estimation of Nonlinear Parameters", J. Soc. Ind. App. Math., Vol. 11, pp. 431-444, June 1963.

15) M.R.Osborne, "Some Special Nonlinear Least Squares Problems", SIAM J. Numer. Anal. Vol. 12, No. 4, pp. 572-592, Sept. 1975.

16) R.Birgenhier, "Parameter Estimation of Multiple Signals", Ph.D Dissertation, University of California, Los Angeles, 1972.

17) P.Bloomfield, Fourier Analysis of Time Series: An Introduction, John Wiley, New York, 1976, Ch:1 and 2.

18) D.W.Tufts and R.Kumaresan, "Improved Spectral Resolution II", Proceedings of IEEE Conf. on Acoust. Speech and Signal Proc., April 1980, pp. 592-597, Denver, CO.

19) D.C.Rife and R.R.Boorstyn, "Multiple Tone Parameter Estimation from Discrete Time Observations", Bell System Tech. Journal, pp. 1389-1410, Nov. 1976.

20) A.V.Oppenheim and R.W.Schafer, Digital Signal Processing, Prentice-Hall, Englewood Cliffs, NJ, 1975.

21) R.N.McDonough and W.H.Huggins, "Best Least-Squares Representation of Signals by Exponentials", IEEE Trans. Auto. Cont., p. 408, Aug. 1968.

22) F.C.Schweppe, "Sensor-Array Processing for Multiple-Signal Sources", IEEE Trans. on Info. Theo. Vol. IT-14, No. 2, Mar. 1968.

23) A.Van den Boss, "A Class of Small Sample Nonlinear Least Squares Problems", Automatica, Vol. 16, pp. 487-490, 1980.

24) K.Steiglitz and L.E.McBride, "A Technique for Identification of Linear Systems", IEEE Trans. on Auto. Cont. Cont. AC-10, Oct. 1965.

25) R.E.Kalman, "Design of a Self-Optimizing Control System", Trans. ASME, Vol. 80, pp. 468-478, Feb. 1958.

26) K.Steiglitz, "On the Simultaneous Estimation of Poles and Zeros in Speech Analysis", IEEE Trans. on Acous., Speech and Signal Processing, Vol. 25, pp. 229-234, 1977.

27) W.J.Done et al. "Evaluation of the Steiglitz Algorithm for Estimating the Parameters of an ARMA Process", Proc. ICASSP 1979.

28) S.M.Kay, "Accurate Frequency Estimate at Low SNR", Submitted to IEEE Trans. on Acoust., Speech and Signal Processing, Aug. 1982.

29) P.Stoica et al. "The Steiglitz-McBride Algorithm-Convergence Analysis etc.", IEEE Trans. Auto. Cont. Vol. 26, June 1981.

30) A.G.Evans and R.Fischl, "Optimal Least Squares Time-Domain Synthesis of Recursive Digital Filters", IEEE Trans. on Audio and Electroacoustics, Vol. AU-21, No. 1, pp. 61-65, Feb. 1973.

31) R.Prony, "Essai Experimental et Analytique etc.", L'ecole Polytechnique, Paris, 1 Cahier 2, pp. 24-76, 1795.

32) F.B.Hildebrand, Introduction to Numerical Analysis, McGraw-Hill, New York, 1956, Ch. 9.

33) C.Runge and H.Konig, "Vorlesungen Uber Numerische Rechnen", Vol. 11 of Die Grundlehren der Mathemat. WissenSchaften, Berlin, Springer, 1924.

34) M.L.Van Blaricum, "A Bibliography on Prony's Method", An Unpublished List of References.

35) M.L.Van Blaricum, "Investigation of Procedures for Automatic Resonance Extraction From Noisy Transient Electromagnetics Data Volumes, I, II, and III", Final Report: ONR Contract: N00014-80-c-0299.

36) T.Y.Young and W.H.Huggins, "Representation and Analysis of Signals. Part XIII. Discrete Orthonormal Exponentials", John Hopkins University, Dept. of Elec. Eng. Tech. Rept. Nov. 1962.

37: W.H.Huggins, "Representation and Analysis of Signals. Part SVII. Reprints on Signal Theory-A Final Report", John Hopkins University, Dept. of Elec. Eng. Tech. Rpt., Nov. 1963.

38) R.N.McDonough, "Representation and Analysis of Signals. Part XV. Matched Exponents for the Representation of Signals", John Hopkins University, Dept. of Elec. Eng. Tech. Rept. April 1963.

39) J.I.Makhoul, "Linear Prediction: A Tutorial Review", Proc. of the IEEE, Vol. 63, pp. 561-580, April 1975.

40) M.Morf et al. "Efficient Solution of Covariance Equations for Linear Prediction", IEEE Trans. on Acou., Speech and Sig. Processing, Vol. ASSP-25, No. 5, Oct. 1977.

41) S.L.Marple, "A New Autoregressive Spectrum Analysis Algorithm", IEEE Trans. on Acous. Speech and Sign. Proc., pp. 442, Aug. 1980.

42) F.R.Spitznogle and A.H.Quazi, "Representation and Analysis of Time-Limited Signals Using a Complex Exponential Algorithm", J. Acoust. Soc. Am., Vol. 47, No. 5, (part 1), 1150-1155, May 1970.

43) J.D.Markel, "The Prony Method and its Application to Speech Analysis", J. Acoust. Soc. Am., Vol. 49, Part 1, p. 105 (A), Jan. 1971.

44) A.E.Siegman and H.Y.Miller, "Unstable Optical Resonator Loss Calculations Using the Prony Method", Applied Optics, Vol. 9, No. 12, pp. 2729-2736, Dec. 1970.

45) T.J.Ulrych and R.W.Clayton, "Time Series Modelling and Maximum Entropy", Physics of the Earth and Planetary Interiors, Vol. 12, pp. 188-200, August 1976.

46) D.N.Swingler, "A Comparison Between Burg's Maximum Entropy Method and Nonrecursive Technique for the Spectral Analysis of Deterministic Signals", J. of Geophy, Res., pp. 679-685, February, 1979.

47) L.B.Jackson et al., "Frequency Estimation by Linear Prediction", Proc. of the Intl. Conf. on Acous., Speech and Signal Proc., pp. 352-356, 1979.

48) M.L.Van Blaricum, "Difficulties Present in Algorithms For Determining the Rank and Proper Poles with Prony's Method", Proc. of the Workshop on Spectral Estimation, Rome Air Development Center, 1979, pp. 147-155.

49) S.W.Lang and J.H.McClellan, "Frequency Estimation with Maximum Entropy Spectral Estimators", IEEE Trans. on Acous., Speech and Signal Processing, Vol. ASSP-28, No. 6, pp. 716-723, Dec. 1980.

50) R.Kumaresan and D.W.Tufts, "Data Adaptive Principal Component Signal Processing", in Proc. of The 19th IEEE Intl. Conf. on Decision and Control, Dec. 10-12, 1980, Albuquerque, NM.

51) V.C.Klemma and A.J.Laub, "The Singular Value Decomposition: Its Computation and Some Applications", IEEE Trans. on Auto. Con., Vol. AC-25, No. 2, April 1980.

52) R.Kumaresan and D.W.Tufts, "Singular Value Decomposition and Spectral Analysis", Proceedings of the IEEE Sponsored First Workshop on Spectral Estimation, August 1981, Hamilton, Ontario, pp. 6.4.1-6.4.12.

53) D.W.Tufts and R.Kumaresan, "Singular Value Decomposition and Frequency Estimation by Linear Prediction", IEEE Trans. on Acou. Speech and Signal Processing, August 1982.

54) D.W.Tufts and R.Kumaresan, "Frequency Estimation of Multiple Sinusoids: Making Linear Prediction Perform Like Maximum Likelihood", To appear in the Proc. of the IEEE, Special Issue on Spectral Estimation, Sept. 1982.

55) R.Kumaresan and D.W.Tufts, "Estimating the Parameters of Exponentially Damped Sinusoidal Signals and Pole-Zero Modelling", To appear in the IEEE Trans. on Acoust., Speech and Signal Proc.

56) D.W.Tufts, R.Kumaresan and I. Kirsteins, "Data Adaptive Signal Estimation by Singular Value Decomposition of a Data Matrix", Proceedings of the IEEE, (Proc.Lett), Vol. 70, No. 7, July 1982.

57) T.L.Henderson, "Matrix Methods for Determining System Poles from Transient Response", Tech. Rept., University of Kentucky, Louisville, Kentucky, May 1980.

58) J.P.Burg, "Maximum Entropy Spectral Analysis", Ph.D. Dissertation, Stanford University, Stanford, CA 1975.

59) A.H.Nuttall, "Spectral Analysis of a Univariate Process with Bad Data Points via Maximum Entropy and Linear Predictive Techniques", in NUSC Scientific and Engineering Studies, Spectral Estimation, Naval Underwater Systems Center (NUSC), New London, CT. March 1976.

60) V.F.Pisarenko, "Retrieval of Harmonics From a Covariance Function", Geophysics J.R.Astron. Soc., 33, pp. 247-266.

61) C.L.Lawson and R.J.Hanson, Solving Least Squares Problems, Prentice-Hall, Englewood Cliffs, NJ 1974.

62) R.R.Hocking and R.N.Leslie, "Selection of the Best Subset in Regression Analysis", Technometrics, Vol. 9, pp. 531-540, 1967.

63) J.R.Khorrami, "Study of Digital Filter Design and Frequency Estimation by Complex Subset Selection", M.S. Thesis, University of Rhode Island, 1979.

64) A.Cantoni and L.C.Godara, "Resolving the Directions of Sources in a Correlated Field Incident on an Array", J. of Acoust. Soc. Am., 67(4), April 1980, pp. 1247-1255.

65) G.Bienvenu and L.Kopp, "Adaptivity to Background Noise Spatial Coherence for High Resolution Passive Methods", Proc. of Intl. Conf. on Acoust., Speech and Signal Processing, April 1980, pp. 307-310.

66) W.F.Gabriel, "Spectral Analysis and Adaptive Array Superresolution Techniques", Proc. of the IEEE, Vol. 68, No. 6, pp. 654-666, June 1980.

67) R.Schmidt, "Multiple Emitter Location and Signal Parameter Estimation," Proc. RADC Spectral Estimation Workshop, pp. 243-258, Rome, NY 1979.

68) S.S.Reddi, "Multiple Source Location-A Digital Approach", IEEE Trans. on Aero. and Elec. Syst., Vol. AES-15, No., pp. 95-105.

69) J.E.Evans et al., "Application of Advanced Signal Processing Techniques to Angle of Arrival Estimation in ATC Navigation and Surveillance System", Tech Rept. 582, Lincoln Laboratories, Mass. Inst. Tech., Lexington, MA, 1982.

70) A.J.Berni, "Target Identification by Natural Resonance Estimation", IEEE Trans. on Aero. and Elec. Syst., AES-11, No. 2, March 1975.

71) D.H.Johnson and S.R.DeGraaf, "Improving the Resolution of Bearing in Passive Sonar Arrays by Eigenvalue Analysis", IEEE Trans. Acoust., Speech and Signal Proc., Vol. ASSP-30, No. 4, pp. 638-647, August 1982.

72) N.L.Owsley, "Adaptive Data Orthogonalization", Proc. Intl. Conf. on Acoust., Speech and Signal Proc. 1978, pp. 109-112.

73) W.S.Liggett, "Passive Sonar: Fitting Models to Multiple Time Series", in Signal Processing, Eds. J.W.R.Griffiths et al., Academic Press, New York, 1973.

74) W.J.Bordelon, "Complementarity of the Reddi Method of Source Direction Estimation with Those of Pisarenko and Cantoni and Godara I", J. Acoust. Soc. Am., 69, pp. 1355-1359, 1981.

75) J.A.Cadzow and T.P.Bronez, "An Algebraic Approach to Superresolution Adaptive Array Processing", Proc. of the Intl. Conf. on Acoust. Speech and Signal Processing, Atlanta, GA, April 1981, pp. 302-305.

76) J.N.Holt and R.J.Antill, "Determining the Number of Terms in a Prony Algorithm Exponential Fit", Math. Biosciences, 36, 319-332, 1979.

77) C.Gueguen et al., "Factorial Linear Modelling, Algorithm and Application", Proc. of the Intl. Conf. on Acoust. Speech and Signal Proc. April 1980, pp. 618-621.

78) C.Lanczos, Applied Analysis, Prentice-Hall, Englewood Cliffs, NJ, 1956, pp. 267-272.

79) G.Strang, Linear Algebra and Its Applications, 2nd Ed., Academic Press, New York, 1976.

80) O.N.Strand, "Theory and Methods Related to the Singular-Function and Landweber's Iteration etc.", Soc. Ind. Appl. Math. J. Numer. Anal., Vol. 11, No. 4, pp. 798-825, September 1974.

81) J.I.Makhoul, "On the Eigenvectors of Symmetric Toeplitz Matrices", IEEE Trans. on ASSP, Vol. ASSP-29, No. 4, pp. 868-872.

82) N.L.Owsley, "A Recent Trend in Adaptive Signal Processing for Sensor Arrays: Constrained Adaptation", In Signal Processing, Eds. J.W.R.Griffiths et al. Academic Press, New York, 1976.

83) C.Eckart and G.Young, "The Approximation of One Matrix by Another of Lower Rank", Psychometrika, 1(1936), pp. 211-218.

84) G.H.Golub and C.Reinsch, "Singular Value Decomposition and Least Squares Solutions", Numer. Math., Vol. 14, pp. 403-420, 1970.

85) H.Hotelling, "Analysis of a Complex of Statistical Variables into Principal Components", J. Educ. Psych., Vol. 24, pp. 417-441, 448-520, 1933.

86) C.R.Rao, "The Use and Interpretation of Principal Component Analysis in Applied Research", Sankhya Ser.A, 26, 329-358, 1964.

87) S.W.Lang and J.H.McLellan, "A Simple Proof of Stability for All-Pole Linear Prediction Models", Proc. of the IEEE (Proc. LETT.), Vol. 67, No. 15, pp. 860-861, May 1979.

88) L.Pakula and S.M.Kay, "Simple Proofs of the Minimum Phase Property of the Prediction-Error Filter", Submitted to the IEEE Trans. on Acoust., Speech and Sig. Proc.

89) R.Kumaresan, "Optimal Frequency Estimation of Sinusoidal Signals in Noise and Applications", M.S. Thesis, Univ. of Rhode Island, 1979.

90) H.H.Lai and J.D.Dyson, "The Determination of the Direction of Arrival of An Interference Field", Radio Science, Vol. 16k No. 3, pp. 365-376, May-June 1981.

91) J.H.Wilkinson, The Algebraic Eigenvalue Problem, Clarendon Press, Oxford, U.K., 1965.

92) R.Kumaresan and D.W.Tufts, "Improved Spectral Resolution III: Efficient Realization", Proc. of the IEEE, (Proc. Lett.) Vol. 68, No. 10, Oct. 1980, pp. 1354-1356.

93) H.J.Larson, Introduction to the Theory of Statistics, John Wiley, New York, 1973.

94) P.R.Krishnaiah, Ed., Multivariate Analysis, North-Holland Publishing Company, 1980.

95) C.R.Rao, "Matrix Approximations and Reductions of Dimensionality in Multivariate Statistical Analysis", in Multivariate Statistical Analysis, Ed. P.R.Krishnaiah, North-Holland Publishing Company, 1980, pp. 3-22.

96) O.S.Halpeny and D.G.Childers, "Composite Wavefront Decomposition via Multi-dimensional Digital Filtering of Array Data", IEEE Trans. on Circuits and Systems, Vol. CAS-22, pp. 552-562, June 1975.

97) N.L.Owsley, "Modal Decomposition of Data Adaptive Spectral Estimates", Presented at the Yale Univ. Workshop on Appl. of Adaptive System Theory, New Haven, CT, 1981.

98) R. Kumaresan and D.W.Tufts, "On Estimating the Angle of Arrivals of Multiple Planewaves", To appear in the Nov. 1982 issue of the IEEE Trans. on Aero. and Elec. Systems.

99) R.Kumaresan and D.W.Tufts, "Accurate Parameter Estimation of Noisy Speech-Like Signals", Proceedings of the Intl. Conf. on Acoust., Speech and Signal Proc. Paris, May 1982.

100) M.L.Van Blaricum, "Problem and Solutions Associated with Prony Method for Processing Transient Data", IEEE Trans. on Antennas and Propagation, Vol. AP-26, No. 1, Jan. 1978.

101) C.Gueguen et al., "Factorial Linear Modelling: Algorithms and Applications", Proc. of the Intl. Conf. on Acoust., Speech and Signal Processing, 1980, pp. 618-621.

102) D.F.Tuttle, Network Synthesis, Vol. 1, John Wiley, New York, 1958.

103) P.R.Aigrain and E.M.Williams, "Synthesis of N-Reactance Networks for Desired Transient Response", J. Appl. Physics, Vol. 20, pp. 597-600, June 1949.

104) D.W.Tufts and R.Kumaresan, "Improved Spectral Resolution", Proc. IEEE (Lett.), Vol. 68, pp. 419-420, Mar. 1980.

OFFICE OF NAVAL RESEARCH
STATISTICS AND PROBABILITY PROGRAM

BASIC DISTRIBUTION LIST
FOR
UNCLASSIFIED TECHNICAL REPORTS

FEBRUARY 1982

	Copies
Statistics and Probability Program (Code 411(SP)) Office of Naval Research Arlington, VA 22217	3
Defense Technical Information Center Cameron Station Alexandria, VA 22314	12
Commanding Officer Office of Naval Research Eastern/Central Regional Office Attn: Director for Science Barnes Building 495 Summer Street Boston, MA 02210	1
Commanding Officer Office of Naval Research Western Regional Office Attn: Dr. Richard Lau 1030 East Green Street Pasadena, CA 91101	1
U. S. ONR Liaison Office - Far East Attn: Scientific Director APO San Francisco 96503	1
Applied Mathematics Laboratory David Taylor Naval Ship Research and Development Center Attn: Mr. G. H. Gleissner Bethesda, Maryland 20084	1
Commandant of the Marine Coprs (Code AX) Attn: Dr. A. L. Slafkosky Scientific Advisor Washington, DC 20380	1
Navy Library National Space Technology Laboratory Attn: Navy Librarian Bay St. Louis, MS 39522	1
U. S. Army Research Office P.O. Box 12211 Attn: Dr. J. Chandra Research Triangle Park, NC 27706	1
Director National Security Agency Attn: R51, Dr. Maar Fort Meade, MD 20755	1
ATAA-SL, Library U.S. Army TRADOC Systems Analysis Activity Department of the Army White Sands Missile Range, NM 88002	1
ARI Field Unit-USAREUR Attn: Library c/o ODCSPER HQ USAEREUR & 7th Army APO New York 09403	1
Library, Code 1424 Naval Postgraduate School Monterey, CA 93940	1
Technical Information Division Naval Research Laboratory Washington, DC 20375	1
OASD (I&L), Pentagon Attn: Mr. Charles S. Smith Washington, DC 20301	1

Copies

Director
AMSAA
Attn: DRXSY-MP, H. Cohen
Aberdeen Proving Ground, MD 21005 1

Dr. Gerhard Heiche
Naval Air Systems Command (NAIR 03)
Jefferson Plaza No. 1
Arlington, VA 20360 1

Dr. Barbara Bailar
Associate Director, Statistical Standards
Bureau of Census
Washington, DC 20233 1

Leon Slavin
Naval Sea Systems Command (NSEA 05H)
Crystal Mall #4, Rm. 129
Washington, DC 20036 1

B. E. Clark
RR #2, Box 647-B
Graham, NC 27253 1

Naval Underwater Systems Center
Attn: Dr. Derrill J. Bordelon
Code 601
Newport, Rhode Island 02840 1

Naval Coastal Systems Center
Code 741
Attn: Mr. C. M. Bennett
Panama City, FL 32401 1

Naval Electronic Systems Command (NELEX 612)
Attn: John Schuster
National Center No. 1
Arlington, VA 20360 1

Defense Logistics Studies Information Exchange
Army Logistics Management Center
Attn: Mr. J. Dowling
Fort Lee, VA 23801 1

Copies

Reliability Analysis Center (RAC)
RADC/RBRAC
Attn: I. L. Krulac
Data Coordinator/
Government Programs
Griffiss AFB, New York 13441 1

Technical Library
Naval Ordnance Station
Indian Head, MD 20640 1

Library
Naval Ocean Systems Center
San Diego, CA 92152 1

Technical Library
Bureau of Naval Personnel
Department of the Navy
Washington, DC 20370 1

Mr. Dan Leonard
Code 8105
Naval Ocean Systems Center
San Diego, CA 92152 1

Dr. Alan F. Petty
Code 7930
Naval Research Laboratory
Washington, DC 20375 1

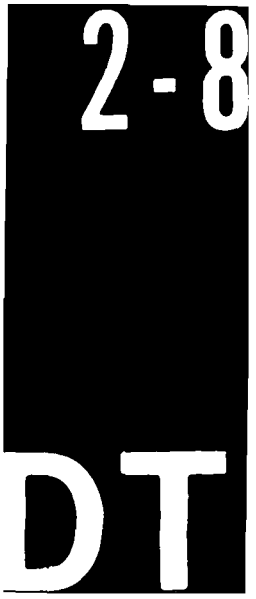
Dr. M. J. Fischer
Defense Communications Agency
Defense Communications Engineering Center
1860 Wiehle Avenue
Reston, VA 22090 1

Mr. Jim Gates
Code 9211
Fleet Material Support Office
U. S. Navy Supply Center
Mechanicsburg, PA 17055 1

Mr. Ted Tupper
Code M-311C
Military Sealift Command
Department of the Navy
Washington, DC 20390 1

Copies

Mr. F. R. Del Priori
Code 224
Operational Test and Evaluation
Force (OPTEVFOR)
Norfolk, VA 23511 1


2-8
DT